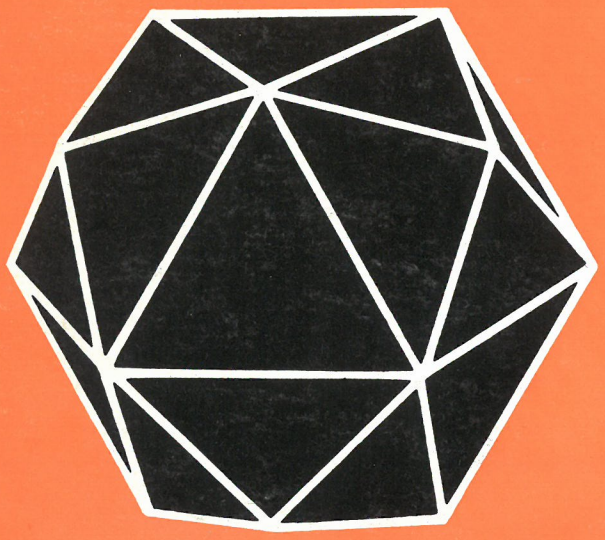
Quaderni di informatica

Rassegna di tecnologia ed applicazioni degli elaboratori elettronici

a cura del Centro Studi Informatica e Automazione Bull Italia



ARCHIVIO STORICO
ASSOCIAZIONE POZZO DI MIELE
FPDM-43
RQDI-134

36
Anno XV - Numero 2 - 1989

FPDM - 43

Worldwide
Information
Systems

Bull

Quaderni di informatica

Rassegna di tecnologia ed applicazioni degli elaboratori elettronici
Anno XV - Numero 2 - 1989

Sommario

L'informatica nell'ufficio: una panoramica

Najah Naffah, uno dei maggiori esperti di applicazione dell'informatica al lavoro d'ufficio, riassume qui i concetti di base e gli orientamenti del settore, con particolare riguardo alle attività che riguardano i manager.

N. NAFFAH

Metodi per l'accesso ad archivi di testi

Il problema di reperire informazioni distribuite in ampie raccolte di testi è di ovvio e crescente interesse. Per risolvere il problema sono stati ideati opportuni approcci metodologici.

L. FELICIAN

Computer e inquinamento elettromagnetico

Le apparecchiature elettroniche generano, durante il loro funzionamento, disturbi di natura elettromagnetica che possono interferire su altre apparecchiature. L'articolo definisce i termini del problema e i modi per contenere entro valori accettabili questa poco nota, ma non per ciò meno importante, forma di inquinamento ambientale.

M. MARTINENGO

Conversazioni col computer

È possibile che un computer sostenga una conversazione con un interlocutore umano riuscendo a mascherare la propria natura? È questa la sfida di "Eliza", uno dei più famosi e provocatori programmi della storia dell'informatica.

F. FILIPPAZZI

Biblioteca dei "Quaderni di Informatica"

Tom Conlon: MicroProlog - Verso la programmazione descrittiva con un linguaggio di 5ª generazione.

Direttore Responsabile
Franco Filippazzi

Comitato di Redazione
Antonio Cicu, Carlo Falcetti, Paolo Lupo, Mario Nobile, Giulio Occhini, Fulvia Sala

Segreteria di Redazione
Elena Pighin

Redazione
Bull Italia
Centro Studi Informatica e Automazione
Via G.M. Vida, 11 - 20127 Milano

Stampa
tipolito Maggioni srl

Autorizzazione del Tribunale di Milano
n. 93 del 20 Marzo 1974

«Quaderni di Informatica» ospita articoli e contributi di varia provenienza. La responsabilità delle opinioni espresse rimane agli Autori.

La riproduzione di articoli della rivista, o di parte di essi, è consentita solo citando la fonte e l'autore.

L'informatica nell'ufficio: una panoramica

NAJAH NAFFAH

Gruppo Bull
Ricerca e Sviluppo
Parigi

1. Introduzione

La tecnologia informatica ebbe la sua prima vera risonanza in tutto il mondo all'inizio degli anni 60. Anche la microelettronica cominciò ad essere conosciuta dal grande pubblico attorno allo stesso periodo, grazie anche alla commercializzazione delle radio a transistor. I servizi di telecomunicazione, come il telefono ed il telegrafo, erano già largamente diffusi nel diciannovesimo secolo. Anche se queste tre tecnologie possono avere avuto origini diverse ed essere apparentemente rivolte a diversi settori di attività, in realtà esse hanno un denominatore comune: la materia prima su cui esse operano è l'informazione. Allo stato attuale, esse risultano saldate insieme in un gruppo unificato di metodologie e tecniche conosciute con il nome di tecnologia dell'informazione, o informatica. Il presente documento esamina come l'informatica venga applicata al lavoro d'ufficio e, in particolar modo, alle attività svolte dai manager. Nel seguito non verrà descritto in dettaglio come funziona questa tecnologia, ma si porrà l'accento sulla descrizione delle principali scelte messe a disposizione dei manager da parte della tecnologia informatica in termini di servizi, funzioni e tipi di sistemi disponibili.

2. Sistemi d'ufficio computerizzati

I computer che costituiscono il cuore di tutti i più importanti sistemi della nuova tecnologia d'ufficio, sono in grado di memorizzare, analizzare, trattare e trasmettere una grande varietà di forme di informazione: testi, grafici, immagini, voci, manoscritti, ecc.

La versatilità dei sistemi computerizzati deriva dal fatto che lo stesso hardware può essere usato per eseguire una molteplicità di applicazioni, aggiungendo semplicemente ad esso nuovo software.

All'inizio, per disporre di una sufficiente potenza di calcolo, bisognava ricorrere ai grandi elaboratori del tipo maiframe installati in ambienti con determinati requisiti, che di solito venivano usati per applicazioni DP (Data Processing), ossia di tipo gestionale.

Tali sistemi erano impiegati principalmente per attività regolate da tutta una serie di routine chiaramente definite per poter operare su grandi volumi di dati, strutturati in flussi organizzati. Il tipico lavoro di Data Processing consiste nel gestire contabilità varie come, per esempio, elaborazione degli stipendi, manutenzione ed aggiornamento di record costituenti gli archivi dei clienti, delle vendite o dei fornitori.

Nei primi sistemi DP, il lavoro doveva essere trasportato fisicamente nel centro di calcolo, dove esso veniva elaborato a lotti (batch).
Gradualmente, però, grazie alla introduzione di particolari unità, i terminali, furono sviluppati collegamenti esterni in grado di assicurare una trasmissione diretta e on-line fra il mainframe e i terminali stessi.

Negli anni '70, furono introdotti i minicomputer, sistemi più piccoli, più robusti e più economici dei precedenti, i quali, dopo una fase di perfezionamento, riuscirono a funzionare correttamente in ambienti normali come, per esempio, gli uffici. Questo processo di decentralizzazione della capacità di elaborazione subì una rapida accelerazione negli anni '70 quando la microelettronica permise di costruire i microcomputer molto più piccoli ed economici, ma potenti e versatili.

Divenne così economicamente e tecnicamente fattibile disporre presso il proprio ufficio (non importa se quello di un manager, o di un segretario, o di un impiegato, o di un professionista) di sistemi computerizzati molto sofisticati.

Tra sistemi DP e sistemi d'ufficio esiste un'ampia sovrapposizione, ma è bene notare alcune caratteristiche cruciali che li distinguono e che derivano più dalla natura del lavoro d'ufficio che dalla tecnologia impiegata; elementi simili di hardware e di software possono, infatti, essere riuniti in modi diversi in funzione delle varie esigenze applicative.

1) Il lavoro DP è altamente strutturato, routinario e prevedibile. Il lavoro d'ufficio tende invece ad essere meno strutturato, più variabile e imprevedibile.
2) Gli utenti dei sistemi DP sono generalmente tecnici con una grossa esperienza professionale

acquisita attraverso la frequenza di corsi altamente specialistici. Gli utenti d'ufficio, siano essi impiegati o alti dirigenti, hanno invece una formazione tecnica meno specializzata, ed in effetti ne richiedono meno. Gli utenti d'ufficio necessitano quindi di sistemi che siano molto più facili da capire e da usare che non i complessi sistemi accettati dagli utenti DP.

3) I sistemi DP eseguono una grande quantità di lavoro per conto dei vari uffici aziendali, come per esempio la contabilità degli ordini, ma sono stati generalmente considerati servizi di supporto, alquanto lontani dalle operazioni quotidiane. La nuova tecnologia d'ufficio è visibile nella normale attività lavorativa, ed il suo successo o insuccesso provoca un impatto più immediato sulla organizzazione aziendale e la sua efficienza.

4) Il personale tecnico, responsabile della gestione dei centri DP, è stato generalmente rappresentato da specialisti di settori completamente diversi dal resto della organizzazione aziendale.
Il personale incaricato della gestione della nuova tecnologia d'ufficio deve, invece, risultare più integrato nella gestione generale dell'organizzazione.

3. La nuova tecnologia d'ufficio

Il lavoro d'ufficio che può essere gestito con sistemi computerizzati, si può far rientrare nelle seguenti categorie:

1) *Strumenti per professionisti.* Strumenti in grado di supportare compiti specifici, come, per esempio, i fogli elettronici che sono semplici ma efficaci nella gestione finanziaria, e i sistemi applicativi destinati ad aiutare architetti, progettisti di veicoli, medici, ecc...

2) *Documentazione.* Creazione di testi stampati e materiale grafico.

3) *Archiviazione/consultazione.* Conservazione di informazioni che possono successivamente essere consultate per estrarre, analizzare o distribuire le parti che interessano.

4) *Trasmissione.* Trasmissione di documenti e di informazioni tra sedi diverse.

La produzione dell'industria impegnata nel settore dei sistemi d'ufficio riguarda le seguenti categorie:

1) *Personal computer.* Microcomputer che offrono risorse di elaborazione di natura generica, adatti per applicazioni relativamente piccole, qualche volta mancanti di qualsiasi prestazione per la trasmissione dei dati.

2) *Sistemi Integrati per gruppi di lavoro.* Forniscono ad un gruppo di lavoro formato da 2-10 persone, le prestazioni necessarie per: a) produrre documenti; b) utilizzare tools vari; c) archiviare/consultare informazioni; d) accedere ad altre prestazioni di elaborazione di solito disponibili all'interno della rete di comunicazione.

3) *Sistemi Dipartimentali.* Sono basati su elaboratori di tipo mainframe o minicomputer o supermicrocomputer, che agiscono da sistemi di coordinamento centrale. Questi elaboratori hanno il compito di rendere disponibili le prestazioni di elaborazione direttamente agli utenti mediante una apposita rete locale, o a più gruppi di lavoro mediante più reti locali.
La stazione di lavoro (Workstation) è un'unità attraverso la quale un utente può accedere ai servizi computerizzati. Le stazioni di lavoro, destinate principalmente a compiti di gestione, hanno cominciato a fare la loro prima apparizione all'inizio degli anni '80. In questo articolo si fa riferimento principalmente a questo tipo di stazione di lavoro. Nelle pagine seguenti saranno descritte le funzioni che i nuovi sistemi d'ufficio possono svolgere sia in termini di servizi individuali, aventi lo scopo di servire una persona per volta, che di servizi multi-utenza, dove gli utenti devono condividere le risorse di elaborazione.

4. Servizi connessi alla stazione di lavoro individuale

Nel 1979, la ditta statunitense Visicorp cominciò a produrre un particolare strumento software chiamato commercialmente Visicalc, in grado di essere usato su microcomputer. Questo prodotto, noto oggi con il nome tecnico di "foglio elettronico", segnò una svolta nell'uso dei sistemi computerizzati da parte dei manager.

Esso dimostrò, attraverso la sua facilità di impiego, che il computer poteva rappresentare uno strumento efficace nell'assistere un manager a svolgere compiti molto importanti.

Nel 1983, in tutto il mondo erano già esistenti più di un milione di fogli elettronici, disponibili in circa 20 versioni, alcune delle quali più potenti del Visicalc. Dopo aver constatato i vantaggi dei fogli elettronici, molti manager incominciarono ad esplorare altri servizi offerti da questa particolare tecnica di elaborazione personale, come il "word processing" e la gestione di archivi personali.

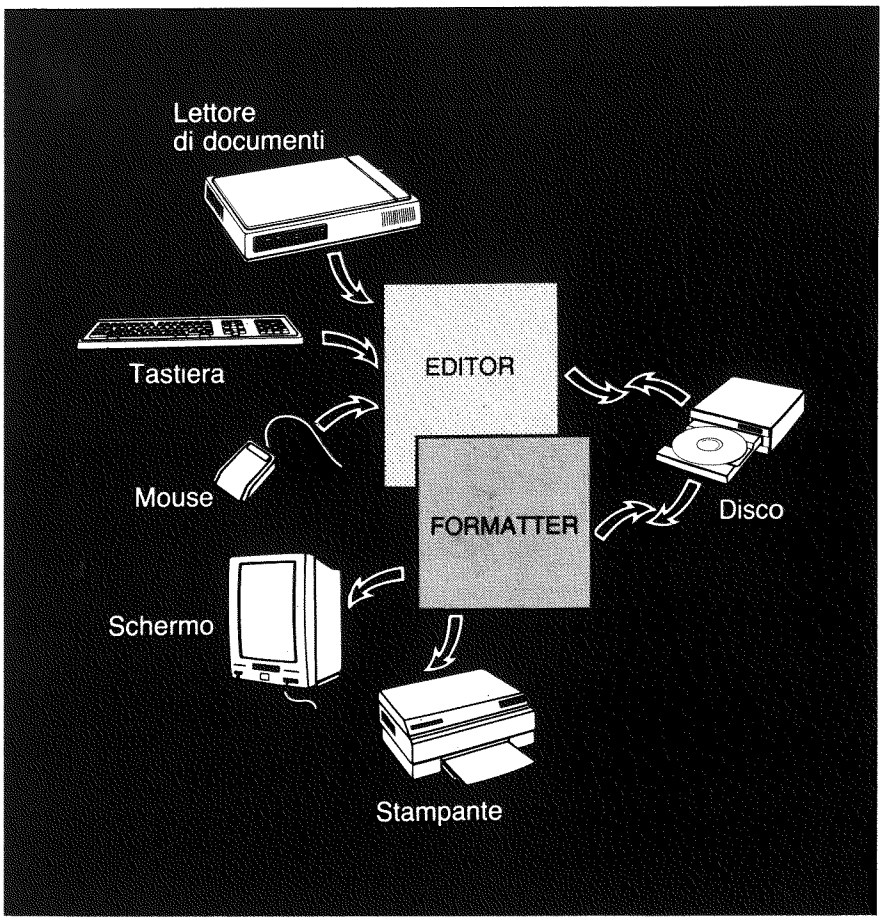
Il software per l'automazione d'ufficio installato su personal computer si può ritenere in grado di soddisfare le esigenze di molti utenti, ma in genere presenta prestazioni meno sofisticate nella produzione di documenti che non i sistemi di editoria professionale progettati per questo uso specifico.

Tuttavia, negli ultimi tre anni, il software per l'automazione d'ufficio installato sui PC ha subito notevoli perfezionamenti.

Word processing e fogli elettronici hanno segnato i primi importanti passi verso servizi più avanzati nel settore della computerizzazione degli uffici.

Il word processing originariamente aveva solo il compito di gestire testi; adesso, però, questa funzione è stata estesa a documenti che contengono anche grafici, diagrammi, tabelle, illustrazioni, fotografie, formule matematiche ed altri simboli speciali. A partire dal 1987 è sorta una nuova industria che si propone di immettere sul mercato prodotti con prestazioni quasi professionali, per sod-

Fig. 1 - Elementi che intervengono nel trattamento di documenti.



disfare le nuove esigenze di "Desk Top Publishing". I fogli elettronici costituiscono un esempio semplificato di modellamento attraverso il software, ovvero la capacità da parte del sistema di gestire regole, equazioni e relazioni che governano i fattori chiave di un problema.

L'utente può usare il modello per scoprire "cosa accadrebbe se" fossero alterati i valori di certe variabili (*What if?*). Il modellamento può diventare estremamente complesso, fino a gestire più variabili e regole di quelle che possono essere gestite su un solo foglio elettronico. Ultimamente ci si stà orientando verso i fogli elettronici "tridimensionali" con i quali è possibile aprire simultaneamente sullo schermo fogli multipli ed eseguire calcoli sui vari elementi che risultano così collegati fra di loro.

Oltre ad occuparsi del trattamento dei documenti e dei fogli elettronici, questo paragrafo descrive altri strumenti di interesse personale per i manager, come l'archiviazione e la consultazione delle informazioni e i miglioramenti dei servizi telefonici.

Nei sistemi computerizzati esistono cinque blocchi fondamentali per trattare e generare un documento (V. Fig. 1).

1) L'input che crea informazioni in forma computerizzata. La tastiera è l'unità di input più comune. Per digitare figure può essere aggiunto al sistema un lettore di documenti (scanner).
2) La memorizzazione dei dati, che ha lo scopo di registrare le informazioni in modo che possano essere modificate durante la prima stesura di un documento o conservate in attesa di essere successivamente consultate.
3) Il software di edizione, che permette all'utente di trattare le informazioni (creazione, cancellazione, copiatura, ecc.).
4) Il software di formattazione, per consentire all'utente di specificare come devono essere presentate le informazioni.
5) L'output che presenta le informazioni. Le due forme di output più comuni sono la stampa su carta e la visualizzazione su video. Il video può anche essere considerato come una finestra per accedere alle informazioni contenute nel supporto di memorizzazione dei dati.

   Questa finestra può rivelarsi troppo piccola per visualizzare tutte le informazioni memorizzate; pertanto può essere spostata verticalmente o orizzontalmente per esplorare tutte le informazioni.

Lo strumento chiave per generare documenti computerizzati è costituito dal software di *edizione* e di *formattazione,* conosciuto anche come "*editor*". Il mercato attuale mette a disposizione questo software con diversi gradi di qualità. Nella valutazione dei sistemi, è importante determinare la facilità e l'efficienza di comunicazione tra utente e computer, nonchè la varietà delle funzioni fornite, compreso il tipo di informazione che può essere trattato.

Gli *editor* recenti sono orientati verso documenti composti o documenti multimedia, contenenti i seguenti tipi di informazioni:

a) Elementi dei testi come: caratteri, parole, sequenze di caratteri (stringhe), frasi, paragrafi, note di piede-pagina ed entità ad alto livello (capitoli, parti, ecc...).

   Tutti questi elementi possono essere definiti in una grammatica formale come si raccomanda nello standard ISO-SGML (Standard Generalized Markup Language). Questa grammatica permette di creare classi di documenti (per esempio, manuale, documento tecnico, lettera...) guidando l'autore nella scelta di un tipo specifico di elementi.

b) Elementi ambientali, di solito associati al materiale del testo, come: gli autori, gli indici, gli allegati e le appendici che si riferiscono all'inizio e alla fine del documento.

c) Informazioni grafiche, come: cerchi, linee, multilinee, linee curve, disegni a manolibera, didascalie delle illustrazioni e così via.

   d) Moduli e tabelle che includono settori, campi e zone, che richiedono di essere trattati.

e) Fotografie o immagini create con punti attraverso un processo di scansione.

f) Voci e suoni che costituiranno un'importante informazione del futuro. Attualmente questi elementi sono usati solo per commenti o annotazioni speciali.

Le funzioni di trattamento comprendono azioni fondamentali come la cancellazione, la copia e il trasferimento, necessarie per tutti i tipi di oggetti anche se la loro implementazione differirà a seconda dei tipi di dati da trattare. Gli editor di testo devono comprendere un gruppo minimo di funzioni dalle quali si possono sviluppare editor estesi. La Fig. 2 illustra un menù, estratto dall'editor Raphael sviluppato dalla Bull, che dà un esempio di quali funzioni debbano essere provvisti i sistemi grafici.

Per quanto riguarda la formattazione, tale processo ha lo scopo di generare un documento rispettando le intenzioni dell'autore in termini di significato e aspetto fisico (sia su video che su carta). L'aspetto non deve andare a scapito del significato, per esempio stampando una tabella su due pagine o confondendo il lettore con troppi stili di caratteri.

I prodotti esistenti attualmente sul mercato mirano a fornire funzioni di formattazione in grado di riunire archivi generati da vari editor (Word, Wordstar, ecc.) e archivi di immagini generati da diversi programmi di disegno (PC Paint, Mac Paint ecc.). Il risultato è di avere un insieme di pagine contenenti oggetti di diverso tipo in formato multicolonna.

Page Maker, e Ventura Publisher sono due esempi di questi prodotti. Altri sistemi sono pienamente integrati e contengono tutte le funzioni di edizione e formattazione (Interleaf, Ploume ecc.).

4.1 Fogli elettronici

Nel passato, contabili e manager erano soliti usare per la loro gestione finanziaria particolari fogli di calcolo che consistevano di una griglia con molte righe e colonne, disegnata su carta con l'ausilio di una penna o di una matita. I fogli elettronici permettono di gestire i dati ed i calcoli richiesti in maniera più efficiente, più completa e più veloce. La tecnica dei fogli elettronici si basa essenzialmente sui seguenti principi:

1) Ogni cella può avere una etichetta (come per esempio il prezzo, l'utile o la percentuale di tassazione), un valore e, se si desidera, vincoli come per esempio un valore massimo o minimo.

2) I vincoli possono collegare una cella alle altre, per esempio attraverso una formula che contenga operazioni aritmetiche (+, -, x, :), logaritmiche, trigonometriche, logiche, operazioni di percentuali e medie, ecc. Per esempio, il valore della cella "utile" potrebbe essere associato ad un vincolo calcolato sottraendo il valore della cella "costi" dal valore della cella "fatturato".

3) Le informazioni fisse relative alla cella etichette, vincoli e interrelazioni formano una mascherina del modello. Quando a particolari celle vengono assegnati nuovi valori, le regole del modello provvedono ad apportare le variazioni dei valori in tutte le celle corrispondenti.

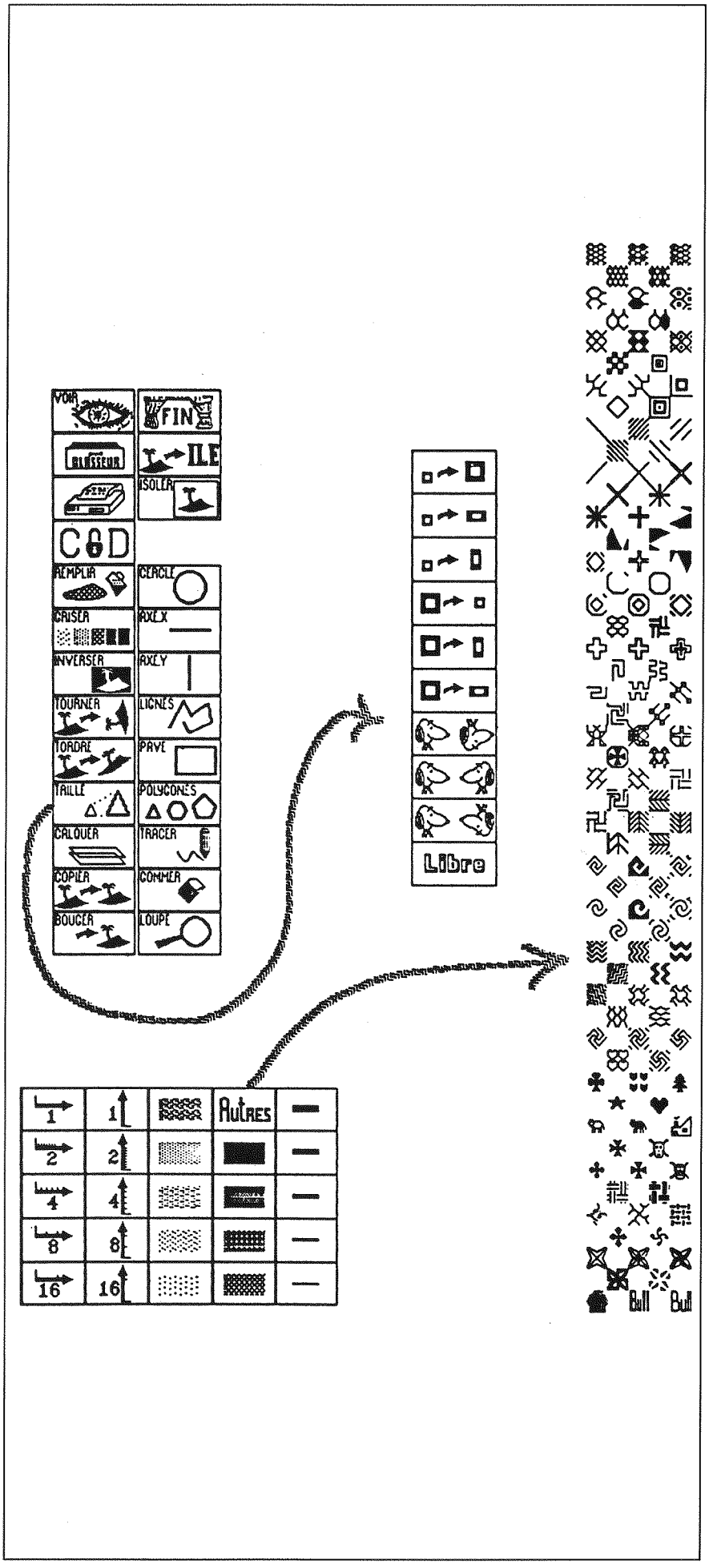
I programmi di fogli elettronici visualizzano solo quella parte di griglia o di tabella che può essere contenuta in una finestra del video. L'indirizzo di ciascuna cella è dato dal suo numero di riga e di colonna all'interno della griglia. Il numero complessivo di colonne e di righe è fisso ed è normalmente compreso tra 64 e 2000 o anche più. Per vedere le celle richieste, l'utente deve far scorrere la finestra. Come descritto più avanti, nel paragrafo riguardante le stazioni di lavoro multifunzionali, alcune stazioni di lavoro permettono all'utente di esaminare contemporaneamente più di una parte di griglia. L'aspetto più tedioso nell'uso dei fogli elettronici consiste nel definire la relazioni fra le celle, fissare i vincoli, assegnare le etichette alle celle, inserire i valori e così via. Tale preparazione, tuttavia, può essere richiesta solo poche volte nel corso di molti anni di utilizzo. Questo compito risulta facilitato se sono disponibili funzioni di edizione che aiutano l'utente, per esempio, a copiare una o più celle, trasferire o scambiare celle, colonne o righe.

Una volta preparato il foglio elettronico, quando l'utente modifica alcune variabili o regole o vincoli, il programma provvede ad eseguire rapidamente tutti i calcoli necessari. Nei sistemi più sofisticati, un certo numero di fogli elettronici possono essere interconnessi attraverso particolari celle: se si modifica una cella, saranno automaticamente apportate le necessarie variazioni a tutte le altre. I fogli elettronici, oltre ad essere in grado di gestire modelli più completi che non le tradizionali tecniche basate su carta e matita, risultano anche molto più veloci e precisi.

I programmi dei fogli elettronici sono usati in svariatissime applicazioni, come previsione delle vendite, pianificazione dei budget, conversione di valute, relazione di bilancio. Attualmente, ci sono più di 20 prodotti utilizzabili sui PC.

I più conosciuti sono: Lotus 1-2-3, Excel, Quattro, ecc.

Fig. 2 - Esempio delle funzioni richieste dai sistemi grafici (in figura: menù estratti dall'editor Raphael della Bull).



4.2 Grafica commerciale

I primi sistemi computerizzati presentavano le informazioni principalmente sotto forma di testo e numeri visualizzati su schermi verdi e grigi, o stampati su pile di tabulati.

Tutto questo scoraggiava molti manager che si allontanavano dai computer un pò delusi perchè essi erano pressati dal proprio lavoro e non potevano dedicare molto tempo a leggere documenti scritti in maniera difficile.

Oggigiorno, esistono molti pacchetti di software che gestiscono la "business graphics" e sono in grado di generare una varietà di forme grafiche, come i diagrammi a barre, i grafici a linee, i diagrammi a "torta" (un cerchio diviso in tanti settori).

Queste presentazioni risultano ovviamente ancor più leggibili se si dispone di un video a colori.

Il software della grafica commerciale è spesso collegato alle prestazioni dei fogli elettronici in modo da tradurre automaticamente i risultati in grafici appropriati. Alcuni sistemi dispongono anche di un editor grafico che può essere utilizzato per modificare le figure generate automaticamente. Una adeguata stampante potrà essere usata per stampare su carta, ma è consigliabile disporre anche della attrezzatura per creare le "slide" che potranno così essere usate direttamente in una riunione o ad una conferenza.

La tendenza attuale è quella di presentare figure e immagini e creare inoltre animazioni sul video con programmi ad hoc. Questa prestazione migliora la semantica del messaggio, soprattutto nel caso in cui il manager voglia mostrare un andamento rispetto al tempo.

4.3 Sistemi di gestione delle informazioni personali

Alcune informazioni sono considerate private e appartenenti ad un individuo o a un gruppo di persone. Pertanto, queste informazioni non sono assoggettate alle regole ed alle procedure aziendali e sono trattate a se stanti. Esistono vari sistemi progettati per gestire tali informazioni personali; per esempio, qualche prodotto software per microcomputer aiuta a creare e gestire alcuni flussi di dati come esatto equivalente elettronico delle schede d'archivio. I sistemi possono essere anche programmati con archivi a scadenza, fissando la data in cui si prevede di non usarli più.

Le "agende elettroniche" possono essere riservate ad un solo utente, come prestazione puramente ad personam, oppure ad un gruppo di utenti nella forma di sistema condiviso, in modo da organizzare riunioni tenendo conto degli impegni che ciascuna persona ha già preso. Gli appuntamenti, con questo metodo, sono fissati automaticamente dal sistema.

Gli archivi di informazioni personali sono registrati, aggiornati e consultati allo stesso modo delle informazioni contenute nei data base più grandi descritti più avanti. L'ultima generazione di prodotti in questo settore è costituita dalle pile di schede contenenti informazioni personali, collegate fra di loro secondo una specifica semantica. Questi prodotti sono basati sul concetto di "ipertesto" ossia sull'accesso associativo a informazioni contenute in testi diversi. La Hypercard commercializzata dalla Apple è un esempio di questi sistemi.

4.4 Potenziamento della comunicazione telefonica

I computer possono contribuire a migliorare i servizi telefonici messi a disposizione di ogni singola persona, per esempio, per mezzo di centralini computerizzati (PABX-Private Automatic Branch Exhange) e con l'adozione di microchip inseriti negli apparecchi telefonici. Grazie a questi accorgimenti si hanno le seguenti prestazioni:

1) Riprova della chiamata. Viene ripetuta automaticamente l'operazione di composizione del numero richiesto, ma che risulta occupato, finchè non è attivato il collegamento.
2) Composizione abbreviata. Si usano numeri telefonici di solo due o tre cifre, per numeri telefonici di uso frequente.
3) Dirottamento della chiamata. La chiamata è dirottata automaticamente su un altro numero interno specificato, nel caso che l'utente si rechi presso un altro ufficio.
4) Ricerca. Se il numero chiamato risulta occupato, la chiamata è dirottata su un certo numero di interni finchè non viene attivato il collegamento.

Nei sistemi più sofisticati, questi tipi di servizi possono essere resi disponibili direttamente da una stazione di lavoro. Per esempio, può essere visualizzato su video un archivio di informazioni personali contenente i nomi e i numeri telefonici. L'utente da chiamare può essere indicato per mezzo di un apposito puntatore. I sistemi computerizzati possono anche registrare voci in forma digitale. Questo metodo può essere utilizzato per registrare i messaggi delle chiamate in arrivo, come avviene per le segreterie telefoniche. Il computer può registrare il messaggio in arrivo in modo che possa successivamente essere ascoltato quando l'utente è libero da altri impegni. Ultimamente, si stanno sviluppando anche sistemi in grado di trasformare i messaggi introdotti da tastiera in voci sintetizzate.

4.5 Sistemi esperti

Un significativo sviluppo delle capacità di elaborazione, che cominciò ad avere un certo impatto al di fuori dei laboratori di ricerca negli anni '80, è rappresentato dal "sistema esperto" che contiene le conoscenze di uno o più specialisti in un particolare argomento o settore. All'inizio, i sistemi esperti furono applicati a ristretti settori scientifici, ma attualmente il loro impiego sta per essere esteso a molteplici applicazioni commerciali.

Essi risultano più facili da realizzare quando si richiede di far gestire loro determinate regole piuttosto che affidar loro il compito di prendere decisioni qualitative.

Alcuni sviluppi dei sistemi esperti sono orientati a sostituire gli specialisti con i computer, ma altre ricerche condotte in questo settore, mi-

rano ad utilizzare l'"intelligenza" messa a disposizione dai computer in modo tale da poter essere integrata con l'esperienza umana.

Alcune nuove applicazioni di questa tecnologia riguardano la gestione delle risorse finanziarie, considerata come un foglio elettronico con capacità deduttive, e il *"computer supported cooperative work"*, diretto a migliorare l'efficienza delle riunioni e della gestione delle procedure amministrative.

5. Servizi con stazioni di lavoro cooperanti

Spesso capita che più persone debbano disporre degli stessi servizi computerizzati per scambiarsi messaggi e per accedere a grandi librerie o archivi di informazioni. L'utilizzo comune dei servizi può avvenire a vari livelli, come per esempio entro una direzione o entro un gruppo di progetto che comprende persone di molte direzioni, o entro un edificio, o tra organizzazioni, o entro una particolare disciplina (chimici, architetti, psicologi, ecc.) o entro una specializzazione di mercato (attività di borsa, attività bancaria, attività di vendita al minuto, ecc.).

5.1 Posta elettronica

La posta elettronica, conosciuta anche come CBMS (Computer Based Message System), mette a disposizione degli utenti una varietà di prestazioni per facilitare i contatti di lavoro tra sedi diverse. Nel 1984, l'ente di standardizzazione CCITT definì una serie di norme rivolte a tutti quei servizi internazionali che si occupano di trasferire messaggi tra CBMS privati. Queste direttive aggiornate nel 1988 sono contenute nel documento "X.400 Recommandations".

Le funzioni del CBMS sono:

1) Generare messaggi. La gamma dei tipi di di informazione (testo, grafica, ecc.) usati nei messaggi, dipende dalle stazioni di lavoro disponibili e dalle prestazioni della rete di trasmissione.
2) Inviare i messaggi ai destinatari. Il software che governa la rete, utilizza informazioni circa le sedi dei destinatari e i canali di trasmissione disponibili per consegnare la comunicazione, nel modo più efficiente e affidabile, al corrispondente computer che contiene la "mailbox" (cassetta postale) indicata.
3) Memorizzare la posta in arrivo finchè il destinatario non è pronto a riceverla. Dopo la trasmissione, il messaggio viene archiviato nella mailbox elettronica dalla quale viene estratto dal destinatario nel momento in cui lo ritiene opportuno. Questo procedimento è conosciuto come sistema di "store and forward" (memorizza e inoltra).
4) Consegnare i messaggi. Il destinatario può regolarmente scorrere un indice per sapere cosa c'è dentro la mailbox. La mailbox è in grado di indicare il nome di chi ha spedito i messaggi, il sommario del suo contenuto, e il grado di priorità o urgenza dei messaggi. I messaggi desiderati possono essere visualizzati su video o stampati su carta.

I servizi di posta elettronica possono essere di grande utilità nelle comunicazioni all'interno di organizzazioni o tra organizzazioni diverse grazie ai seguenti vantaggi:

1) *Nessun vincolo di tempo.* Il mittente può comporre e inviare il messaggio quando lo desidera assegnandogli il grado più opportuno di urgenza. Per fare ciò, non occorre che il destinatario sia presente nello stesso momento in cui si invia il messaggio all'altro capo della linea. Questa caratteristica è ovviamente preziosa quando si trasmette tra stati con diversi fusi orari.

2) *Nessuna interruzione.* La prestazione di store and forward comporta il vantaggio che il destinatario non sia interrotto tutte le volte che arriva un messaggio.

3) *Nessuna limitazione geografica.* La posta elettronica estende le potenzialità del telefono di mettere in contatto persone in posti diversi e può instaurare tra gli utenti un sentimento di stretta comunità. La prestazione di conferenza mediante computer, per esempio, permette agli abbonati di partecipare ad un dialogo continuo tramite la registrazione delle proprie idee in un archivio comune. Un partecipante può unirsi alla conferenza in qualsiasi momento, dopo aver letto gli ultimi commenti.

4) *Multidestinazione.* I messaggi possono essere inviati contemporaneamente a molti destinatari. Liste di mailing possono essere memorizzate nel sistema e richiamate, se necessario, per determinate categorie di comunicazione.

5) *Messaggi fonici.* Se il destinatario desidera aggiungere una nota prima di ritrasmetterla al mittente e/o a qualcun altro, egli deve poter registrarla a voce e inviarla al nuovo destinatario. Per fare ciò, il sistema provvederà a convertire la voce in segnali digitali.

6) *Controllo e monitoraggio delle trasmissioni.*
   a. Per messaggi particolarmente importanti, il mittente può, per esempio, segnalare l'urgenza del messaggio facendo lampeggiare una luce sul terminale del destinatario. Esempi di questi messaggi sono i seguenti: richiedere di essere informato quando il messaggio viene consegnato e letto; specificare destinatari alternativi se il destinatario principale non può essere raggiunto; permettere a tutti i destinatari di sapere chi altri hanno ricevuto le copie, o ottenere copie silenti non segnalate agli altri destinatari.
   b. Il destinatario può, per esempio, dirottare tutta la posta in arrivo su un'altra stazione di lavoro se egli dovrà assentarsi dall'ufficio per un certo periodo; o dirottare la posta proveniente da alcuni abbonati su altre stazioni di lavoro per evitare di ricevere troppi messaggi inutili.

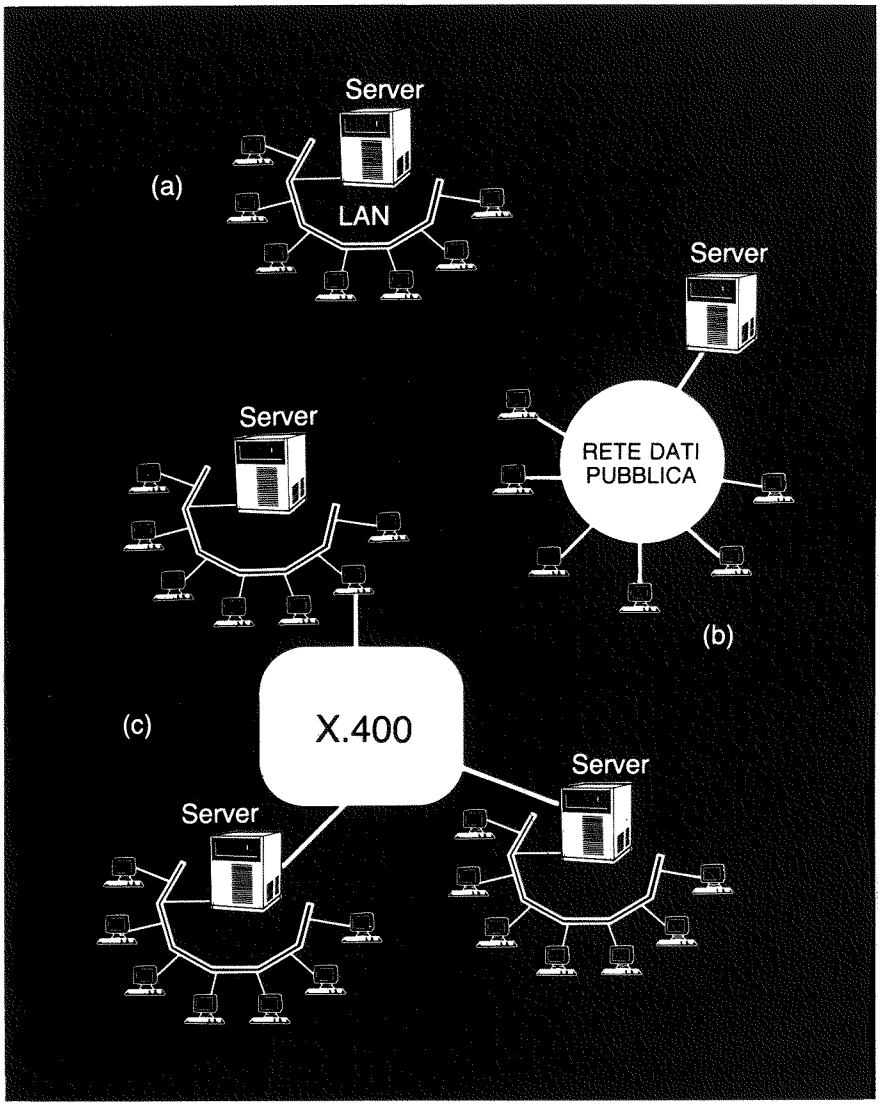


Fig. 3 - Tre classi di soluzioni per la posta elettronica.

7) *Gestione del sistema.* Il sistema richiede di essere amministrato sistematicamente ed efficientemente. Occorre stabilire anche adeguate procedure per assegnare a ciascun abbonato una "password" codificata di identificazione e appropriate infrastrutture di stazioni di lavoro e di rete. L'utilizzo del sistema deve essere monitorato a scopo di rilevamento costi e per ragioni di contabilità. Devono essere aggiornati liste di distribuzione ed altri archivi come quello contenente nomi e attributi di abbonati, in grado, per esempio, di trovare tutti gli altri abbonati residenti in un determinato quartiere della città. Stazioni di lavoro, unità di memorizzazione, stampanti, computer ed altri elementi che fanno parte di un sistema di posta elettronica, possono essere collegati insieme tra di loro in vari modi come illustrato nella Fig. 3, che mostra tre diverse soluzioni:

a. impianto locale con stazioni di lavoro servite da un micro o minicomputer, comprendente mailbox, elenchi e così via;

b. servizio centralizzato su scala nazionale in cui minicomputer e mainframe hanno il compito di coordinare e controllare le stazioni di lavoro collegate per mezzo di una rete, spesso gestita dal servizio pubblico di Poste e Telegrafi o da una rete privata;

c. configurazione distribuita attorno ad un Servizio Trasferimento Messaggi Internazionale: questo è il tipo più recente di servizio definito dalla CCITT e capace di interconnettere sistemi decentralizzati.

5.2 Teleconferenza

La teleconferenza rappresenta un mezzo alternativo alle riunioni faccia a faccia. La conferenza computerizzata, realizzata con i sistemi descritti precedentemente, può essere considerata come una forma di teleconferenza, ma questo termine è di solito usato per quei servizi che permettono agli utenti di interagire fra di loro immediatamente. I due principali tipi di teleconferenza sono:

a. Audioconferenza in cui gli utenti possono parlare fra di loro, ma non possono vedersi. Possono essere forniti dispositivi per scambiare documenti elettronicamente o per trasmettere messaggi di telescrittura usando speciali tavolette di input, che trasmettono i movimenti fatti da una penna. I partecipanti sono riuniti in sale, situate in sedi diverse, dove vengono messi a disposizione

speciali equipaggiamenti per consentire una libera conversazione senza la necessità di dover fare uso di auricolari o microfoni ingombranti.

b. Videoconferenza in cui gli utenti possono vedersi e sentirsi l'un l'altro. Le principali limitazioni tecniche sono quelle di richiedere collegamenti per trasmissioni a larga banda; le animazioni possono richiedere un minimo di 1,5-2 milioni di bit/sec. In sostituzione della conferenza ad animazione completa, è possibile disporre dei servizi di semianimazione che risultano più economici grazie alla loro caratteristica di usare solo 50.000 bit/sec.

   Con questo sistema, viene inviato un quadro statico circa ogni 30 secondi; l'efficacia delle immagini con il sistema della semianimazione non è ovviamente paragonabile a quello della animazione completa. Per le videoconferenze occorre disporre di speciali telecamere, studi, monitor TV ed altro equipaggiamento. Nuovi algoritmi di codifica permettono di scambiare immagini in movimento su canali di 384.000 bit/sec e 56.000 bit/sec.

Le teleconferenze offrono servizi potenzialmente molto interessanti, ma il loro utilizzo è cresciuto molto lentamente.

Ciò è dovuto in parte ai costi che rimangono piuttosto alti se si confrontano con i vantaggi che derivano da questo sistema. Come per altri servizi computerizzati, la microelettronica sta riducendo tali costi, rendendo così la videoconferenza una scelta economicamente valida in molti casi.

Un altro freno alla crescita della teleconferenza è da attribuirsi alla resistenza da parte dei manager ad accettare questo tipo di trasmissione a distanza, perchè le persone riprese dalla telecamera assumerebbero atteggiamenti goffi e innaturali.

Ulteriori esperienze, fatte con riprese in diretta, stanno gradualmente mettendo a disposizione dei manager sistemi che consentono di sfruttare effettivamente la teleconferenza.

5.3 Trasmissione di documenti facsimile

Il sistema Telefax o Facsimile (Fax) serve a scandire un documento, trasmettere la sua immagine e riprodurne una copia sull'unità ricevente. Le prime macchine erano piuttosto lente e generavano copie spesso confuse e illeggibili. Le tecniche digitali computerizzate hanno migliorato velocità e qualità e ridotto i costi. Gli standard CCITT hanno classificato le macchine telefax in quattro categorie. Le categorie 3 e 4 sono le più interessanti per l'ufficio del futuro. I sistemi che rientrano nella categoria 3 inviano la copia di una pagina in meno di un minuto (all'incirca 20 secondi). I sistemi della categoria 4 devono essere altrettanto veloci e non generare errori. Le altre due categorie comprendono principalmente vecchie macchine in grado di trasmettere una pagina del formato A4 in circa sei minuti (categoria 1) o tre minuti (categoria 2).

6. Database aziendali

L'archiviazione, la memorizzazione e il reperimento delle informazioni costituiscono le comuni attività d'ufficio.

In passato, armadi e dossier per archiviare la documentazione erano parte integrante del lavoro d'ufficio. I database computerizzati offrono oggi nuovi mezzi per gestire tali informazioni, riducendo la quantità di carta e di spazio necessaria, accelerando la ricerca per estrarre le informazioni richieste e mettendo in grado i dipendenti di molti uffici di accedere alla stessa fonte di informazioni.

I primi sistemi per la gestione dei database (DBMS), furono sviluppati su sistemi di elaborazione con dati altamente strutturati, come per esempio i record riguardanti la contabilità, gli ordini, le consegne, i clienti, gli inventari, i fornitori e il personale di una azienda. Sono stati usati tre tipi principali di DBMS: i database Gerarchici, di Rete e Relazionali costituiscono l'approccio più recente per ottenere una grande facilità d'uso del sistema, ma comportano controlli di software estremamente complessi. Il database è considerato come un insieme di elementi di dati la cui struttura non è definita al momento della preparazione del database. Le relazioni sono, invece, formate dinamicamente a seconda delle attività in esecuzione. La difficoltà di gestire molteplici relazioni complesse può rendere questo software molto lento.

Per evitare questo inconveniente, sono state messe a punto tecniche per migliorare i componenti hardware da cui dipende la velocità dei sistemi relazionali, come per esempio l'implementazione di alcune routine in firmware, circuiti microelettronici costruiti in modo particolare, o uso di microprocessori per facilitare la ricerca delle informazioni sui dischi.

6.1 Consultazione delle informazioni e linguaggi di interrogazione

Il linguaggio di interrogazione, o ogni altro metodo disponibile agli utenti per ricerche nei database, si basa sulle operazioni di "inquiry" e di "retrieve". Le principali tecniche di ricerca sono qui riassunte.

1) Ricerca con parole-chiave "indexed", in cui gli elementi contenuti negli archivi sono associati a chiavi esplicite di ricerca. I sistemi completi creano liste che identificano il verificarsi di ciascun termine e permettono di fare una ricerca sull'intero testo. I sistemi parzialmente indexed, rappresentati per esempio dalle biblioteche computerizzate, si basano sulla ricerca di attributi, come autore, titolo e soggetto, selezionati da un indexer professionale.
2) La ricerca tramite esemplificazione (Query by Example), realizzata per i database relazionali, presenta all'utente una lista di attributi. L'utente inserisce i valori degli elementi da ricercare. Questo concetto è stato recentemente esteso al paradigma del "box di dialogo" La Fig. 4 mostra un'interfaccia che comincia a diventare comune sugli attuali PC.

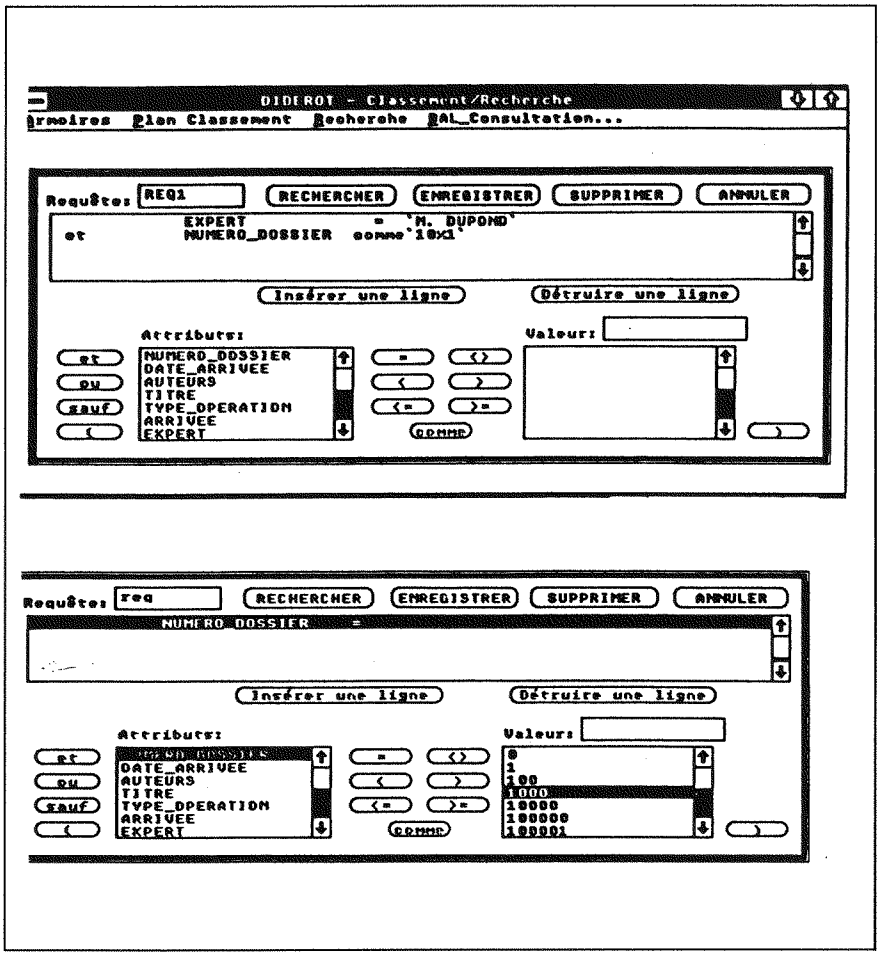


Fig. 4 - Consultazione di basi di dati tramite esempio (Query by example).

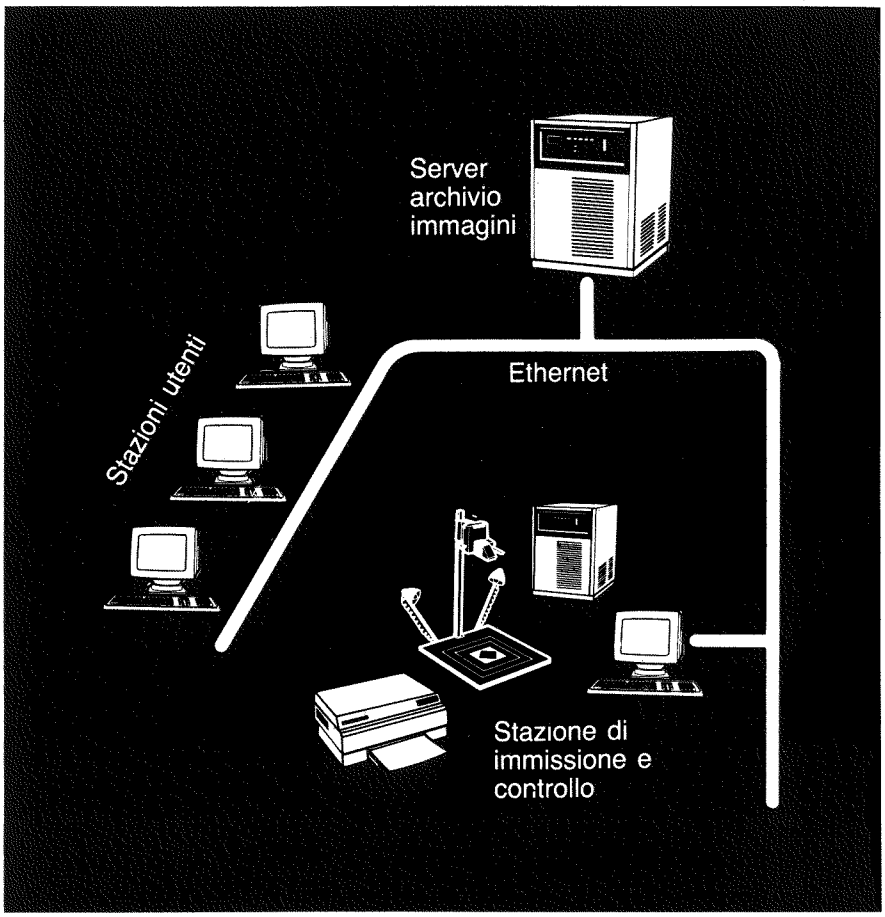


Fig. 5 - Una banca di immagini: principali componenti del sistema SGDM della Bull.

3) Menù che presentano all'utente alcune scelte dalle quali è possibile selezionare o l'informazione richiesta o altri menù che guidano l'utente lungo un percorso verso l'informazione richiesta.

4) Il linguaggio naturale permette all'utente di esprimere richieste usando frasi simili al modo in cui è normalmente espressa la ricerca, come per esempio "Trovami tutti i clienti i cui ordini sono in attesa da più di due settimane". Questi sistemi di norma "capiscono" il linguaggio naturale soltanto entro un determinato contesto e riconoscono soltanto un vocabolario limitato.

6.2 Banche di immagini

I dischi ottici, di cui si è parlato prima in termini di sistemi di informazioni personali, possono essere inseriti dentro unità tipo juke-boxes in modo da disporre di una fonte di informazioni conservate sotto forma di immagini, condivisibile da molti utenti.

Per trasferire i documenti sui dischi si usa un lettore di immagini. Le informazioni riguardanti ciascuna pagina sono introdotte da uno schermo della stazione di lavoro di controllo. È importante ricordare che le informazioni su tali dischi sono memorizzate come immagini di pagine intere (qualche volta in forma compressa).

Esse non possono essere quindi cercate in base a dettagli, come succede in un database per testi.

I principali componenti di un sistema di memorizzazione ottica, con accessi distribuiti su PC, sono mostrati nella Fig. 5. Questo sistema chiamato SGDM è un prototipo prodotto dalla Bull in Francia.

7. Stazioni di lavoro per uffici

Le stazioni di lavoro più avanzate sono costituite da sistemi di piccoli computer che possono funzionare come unità indipendenti o come parti di una rete. Questi computer nacquero negli anni '70 presso il Centro di Ricerca di Palo Alto (PARC) della Xerox, in California, con l'obiettivo di integrare diverse attività e tipi di informazione (testo, grafica, immagini, ecc.) Essi cominciarono ad essere conosciuti nei primi anni '80 quando furono immessi sul mercato modelli come Star, Lisa, Macintosh.

Attualmente, la maggior parte degli utenti di tipo commerciale dispone di stazioni di lavoro costituite da PC. In campo scientifico, troviamo pure stazioni di lavoro che presentano caratteristiche equivalenti, qui di seguito accennate (V. Fig. 6).

1) Visualizzazioni multifinestra in grado di mostrare le diverse applicazioni in atto nella macchina, o diverse parti dello stesso documento o foglio elettronico.

   Per esempio, una finestra potrebbe essere usata per operazioni di edizione, un'altra per visualizzare i contenuti di una mailbox elettronica ed una terza per mostrare una guida telefonica elettronica. Inoltre, le finestre possono essere spostate o rimpicciolite allo scopo di avere più spazio per altre finestre. I "menù" possono essere considerati delle particolari finestre che contengono comandi sotto forma di testo o icone. I "box di dialogo" sono pure visualizzati sotto forma di finestre per acquisire i parametri di un comando complesso o di una interrogazione. Le multifinestre sono di grande importanza per il manager, il cui lavoro può richiedere di seguire contemporaneamente parecchie attività. Esse contribuiscono anche a stabilire una buona comunicazione tra computer e utente.

2) Mezzi di puntamento necessari per usare sistemi multifinestra. Essi sono usati, per esempio, per spostare finestre o altri oggetti sullo schermo o per puntare ad una icona. Le principali unità di puntamento includono i dispositivi seguenti:

   a. Mouse, dispositivo che può essere tenuto nel palmo di una mano, munito di una sfera montata nella base o di un sistema a infrarosso. Il cursore sullo schermo viene mosso nella direzione seguita dal mouse mentre la sua sfera viene fatta rotolare sulla scrivania davanti al video. Il mouse può avere uno o più pulsanti di controllo per essere in grado di svolgere vari compiti.

   b. Schermo sensibile al tocco, in grado di registrare il punto dello schermo che viene toccato. Questo metodo è utile per spostarsi all'interno di menù rappresentati da box di ampie dimensioni, ma risulta difficile da usare nel caso si debba puntare con molta precisione su piccole immagini.

   c. Tavolette, costituite da un supporto sensibile, in grado di registrare la posizione toccata. Sono utili particolarmente in ambito Computer Aided Design (CAD), dove è necessario creare disegni dettagliati.

   d. Penna luminosa, dispositivo elettronico rassomigliante ad una penna, usata per introdurre immagini grafiche, anche questa particolarmente adatta per compiti di CAD.

3) Visualizzazione ad alta risoluzione in grado di mostrare immagini nitide, di solito con caratteri ed immagini neri su sfondo bianco. Le immagini sono visualizzate sullo schermo come aggregazione di minuscoli punti le cui posizioni sono chiamate pixel. Più numerosi sono i pixel di cui si dispone e più alta risulta la risoluzione e la qualità dell'immagine. Schermi contenenti 1024x768 pixel possono riprodurre una pagina ad alta qualità avente il formato A4.

4) Potenti microprocessori aventi prestazioni simili a quelli usati nei micro e minicomputer, tipicamente macchine a 32 bit. Agli ormai classici Intel 386 e Motorola 68020-30 si sono aggiunti nuovi tipi di microprocessori, come componenti di base delle stazioni di lavoro; sono i cosidetti RISC (Reduced Instruction Set Computers), che si prevede negli anni 90 possano raggiungere i 50-200 MIPS.

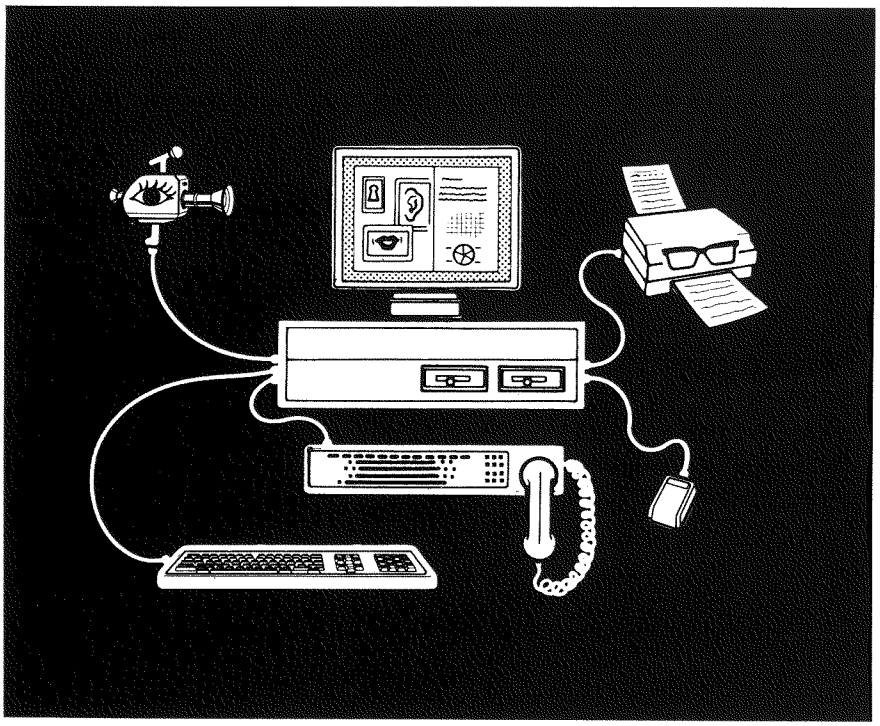


Fig. 6 - Illustrazione del concetto di stazione di lavoro sofisticata.

5) Grande memoria centrale e grande memoria su dischi.
La prima contenente fino a 16 Mbyte, la seconda con capacità da 50 a 200 Mbyte.

8. Reti

Le reti per il lavoro d'ufficio si estendono all'interno di un particolare edificio o gruppo di uffici (distribuzione locale) o su una grande area geografica.

Per i manager, la realizzazione della infrastruttura di telecomunicazione costituisce un obiettivo strategico e tattico.

8.1 Reti per grandi aree

Le reti telefoniche e telex sono classici esempi di Wide Area Networks (WAN). Le prestazioni della rete sono determinate da due principali fattori: la larghezza di banda del mezzo di trasmissione e il metodo usato per inviare i messaggi attraverso il sistema. Il filo di rame delle reti telefoniche offre una piccola larghezza di banda, mentre i cavi coassiali, le fibre, i collegamenti a microonde e via satellite hanno larghezze di bande molto più ampie. I WAN si servono di due principali tipi di commutazione:

- Commutazione di circuito. È la tecnica che realizza il collegamento tra il circuito di rete del dispositivo chiamante e il circuito ricevitore, attraverso una serie di centrali telefoniche per le interconnessioni.
- Commutazione di pacchetto. È la tecnica più usata attualmente nella maggior parte delle nazioni. I dati sono divisi in piccoli pacchetti delimitati da informazioni di controllo, comprendenti una intestazione che specifica la destinazione dei pacchetti. I commutatori della rete inviano ciascun pacchetto al ricevitore attraverso la via più efficiente.

Per assicurarsi che i pacchetti arrivino intatti a destinazione e nel giusto ordine, vengono sottoposti ad appositi controlli. Alle reti a commutazione di pacchetti si applica lo standard X.25.

8.2 Centrali automatiche private

Molte aziende si sono avvalse per molti anni di reti di telecomunicazione distribuite localmente nella forma di comunicazioni telefoniche interne controllate da un centralino privato, costituito generalmente da un pannello di commutazione manuale. Questi impianti sono in via di sostituzione con i PABX (Private Automatic Branch Exchange) computerizzati, in grado di offrire una varietà di servizi per le informazioni d'ufficio, come la telefonia potenziata, la posta elettronica e l'audioconferenza. Uno tra gli svantaggi a lunga scadenza cui vanno incontro i PABX può essere quello che i cablaggi esistenti non abbiano una larghezza di banda sufficiente per convogliare alti volumi di traffico. Un importante vantaggio, tuttavia, è dato dal fatto che può essere installata una rete discretamente potente senza richiedere un ricablaggio rilevante.

8.3 Reti per aree locali

Esistono due principali famiglie di reti per aree locali (LAN):

- reti a larga banda che possono trasmettere molti canali su diverse frequenze, in modo tale da convogliare contemporaneamente un mix di informazioni;
- reti in banda base, con un solo canale che consente la trasmissione di un flusso di dati per volta.

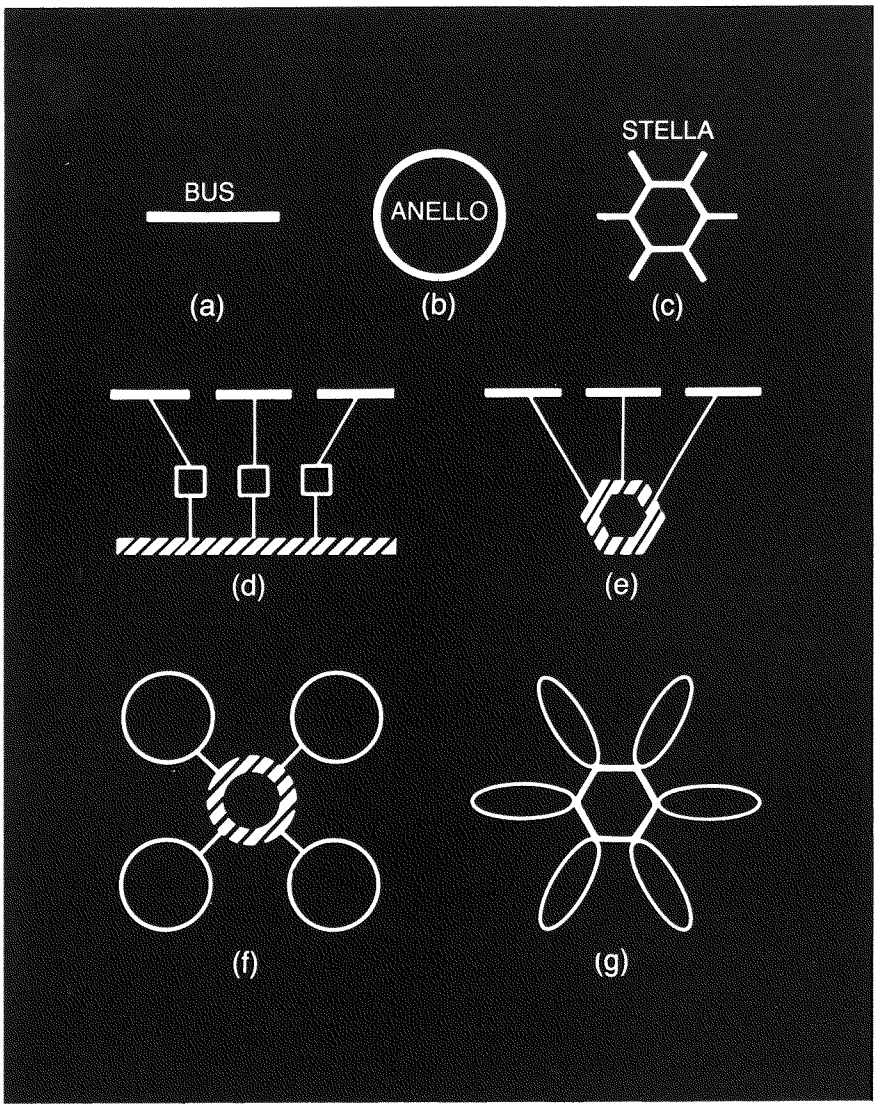


Fig. 7 - Reti locali: a/b/c = topologie di base; d/e/f/g = topologie miste.

Le reti a larga banda possono essere usate per una varietà di scopi, come le trasmissioni TV, le trasmissioni audio multicanale, il controllo ambientale e le reti locali TV via cavo.

Esse servono generalmente un'area geografica più grande che non i sistemi a banda base, limitati di solito a lunghezze di cavo dell'ordine di tre chilometri, anche se LAN indipendenti possono essere collegate fra di loro.

La velocità richiesta per un servizio banda base è normalmente di almeno 1 milione di bit al secondo. I principali tipi di tecniche LAN per la banda base sono i seguenti:

- CSMA-CD (Carrier Sense Multiple Access with Collision Detection), rappresentato dalla rete Ethernet che fu sviluppata nel '75 presso il centro PARC della Xerox e adottato sin da allora dalla maggior parte dei produttori di sistemi per ufficio. In questi sistemi, ciascuna stazione può diffondere messaggi verso altre stazioni, le quali rimangono costantemente in allerta per captare messaggi ad esse indirizzate.
- Ring Empty Slot Systems, sviluppati originariamente presso l'Università di Cambridge, in Gran Bretagna. Sono quei sistemi dove "slot" di dati con lunghezza fissa circolano continuamente nell'anello. Le stazioni di lavoro scaricano le informazioni nelle "slot" vuote o vi raccolgono dati ad esse indirizzati.
- Ring o Bus Token Systems, sono quei sistemi dove le stazioni possono trasmettere soltanto quando sono in possesso di un token, ossia di un codice speciale che circola assieme ad altri dati.

Normative che regolano i vari aspetti delle LAN sono state stabilite da autorità riconosciute in campo internazionale, come IEEE (Institute of Electrical and Electronics Engineers), ECMA (European Computer Manufacturers Association) e ISO.

8.4 Scelta e combinazione di reti

Ciascun tipo di LAN o di PBX per comunicazioni locali, presenta particolari vantaggi e svantaggi. La scelta della tecnica più appropriata per ogni specifica circostanza dipende dai sistemi già installati, dai tipi di informazioni che si vogliono trasmettere, dai costi di collegamento delle unità alla rete e di installazione della rete, dalla prestazione globale, e così via. Per molte aziende, la soluzione più appropriata si può ottenere combinando più tipi di rete. (Fig. 7).

Per esempio, la tendenza attuale è di adottare sistemi a banda base per gruppi di lavoro locali, come tutti gli uffici di un piano di un edificio, per poi collegarli attraverso un sistema a larga banda che serve molti edifici. È anche possibile avere un sistema integrato formato da:

- una o più reti di area locale;
- sistemi di computer individuali, stazioni di lavoro, database e unità di input o output (conosciute come "server" di rete);
- "gateway" che collegano reti private e pubbliche

9. Gli orientamenti

Questo articolo ha posto l'accento su sistemi e servizi ormai operanti in ambito aziendale. Ma l'innovazione tecnologica sta procedendo a ritmo frenetico, con la conseguente continua immissione nel mercato di nuove opzioni.

Le principali tendenze si possono così riassumere:

1) Ulteriore uso di potenti chip per ridurre il costo e le dimensioni delle unità di elaborazione, con incremento della potenza di elaborazione a 10 MIPS entro la fine degli anni '80 e probabilmente a 50-200 MIPS negli anni '90.
2) Miglioramento della comunicazione uomo-computer attraverso nuovi dispositivi di input/output, come sistemi in grado di capire il linguaggio umano e dispositivi di presentazione "amichevoli".
3) Maggiore uso di sistemi basati sulla conoscenza e di documenti multimediali.

Tutto lascia prevedere che l'evoluzione delle tecnologie d'ufficio, già straordinaria nel recente passato, continuerà senza soste anche nel prossimo futuro.

Metodi per l'accesso ad archivi di testi

LEONARDO FELICIAN

Dipartim. di Matematica e Informatica
Università di Udine

1. Introduzione

Questo lavoro affronta il problema delle organizzazioni dati utilizzate per memorizzare i testi, mettendo in luce le differenze con le organizzazioni usate per gli archivi di registrazioni.

Anche se alcuni di questi metodi (organizzazioni sequenziali, liste invertite) si possono applicare parimenti agli archivi testuali, per essi sono state studiate anche soluzioni specifiche, a causa della vasta importanza che il problema della gestione dei testi sta assumendo in contesti diversi: il problema della rappresentazione e della gestione di documenti si affronta nel settore disciplinare attualmente noto con il nome di reperimento dell'informazione (*information retrieval*).

Dopo una introduzione al problema del reperimento di informazioni nei testi, che mette in evidenza le differenze di struttura, tipi di operazioni ed ambiente operativo rispetto alla gestione delle registrazioni negli archivi, vengono esposte le principali tecniche di organizzazione per i testi.

2. Definizione del problema

Il problema di realizzare organizzazioni che consentano un rapido accesso ad informazioni memorizzate in forma testuale (*text retrieval*) ha giustificato un vasto lavoro di ricerca che ha preso le mosse dall'ambiente della consultazione di biblioteche (libri, giornali e riviste) e di sommari di pubblicazioni scientifiche (*abstracts*), estendendosi successivamente entro i confini dell'automazione del lavoro d'ufficio (lettere, studi, relazioni). Nel seguito si parlerà in generale di *documenti* per indicare informazioni che hanno tra le loro caratteristiche la presenza di un testo e possono essere reperiti specificando condizioni sul testo stesso.

Per inquadrare la rilevanza di questo problema bastano poche considerazioni. In ogni ufficio viene trattata una grande quantità di informazioni eterogenee che subiscono processi di distribuzione, consultazione ed archiviazione: recenti stime indicano che soltanto 1/3 di esse rientrano nella categoria strutturata, mentre il resto è costituito da testi, immagini, voce. Tutte queste categorie di informazioni vanno correntemente sotto il nome di *campi lunghi* (*long field*), per evidenziare da un lato l'assenza di struttura, dall'altro l'occupazione di spazio richiesta. Il passaggio dall'una all'altra categoria di informazione richiede un salto in media di un ordine di grandezza nel valore tipico di occupazione di spazio, come risulta dalla Fig. 1.

L'automazione di numerose funzioni connesse con il trattamento delle informazioni non strutturate può ridurre drasticamente i tempi oggi impiegati per la loro gestione ed aumentare in maniera considerevole la produttività nel settore terziario. Le tecniche di archiviazione elettronica, tra l'altro, si integrano in maniera naturale con l'elaborazione dei testi, la posta elettronica e l'elaborazione delle immagini, tutte aree di grande interesse nelle realtà aziendali odierne.

Le differenze salienti tra le organizzazioni in un testo e quelle utilizzate sui dati si riferiscono alla struttura delle informazioni, ai tipi di interrogazione ed ai requisiti operativi.

2.1 Struttura delle informazioni

La basi di dati sono insiemi di dati strutturati in relazione tra loro: invece i testi sono considerati in genere privi di struttura, come una lunga sequenza di caratteri, oppure come dati strutturati in cui le componenti elementari sono le *parole,* sequenze di caratteri separate da un separatore (spazio bianco o segno di interpunzione), articolate in *frasi,* sequenze di parole terminate da particolari segni di interpunzione (punti, punti di domanda, punti esclamativi). A loro volta le frasi sono raggruppate in *paragrafi,* sequenze di frasi terminate da un'indicazione di capoverso.

Da questo punto di vista, non esiste nulla che possa essere assimilato in linea di principio agli attributi di una registrazione in formato fisso: l'unico metodo di accesso possibile in maniera nativa ad un archivio di testi è quindi la scansione sequenziale completa, lenta e solitamente di difficile applicabilità a causa del grande volume delle informazioni testuali.

È evidente la necessità di cercare

Categoria di informazioni	*Occupazione tipica*
Dati (una registrazione)	100 byte
Testi (una pagina in formato A4 dattiloscritta a spazio 2)	2 Kbyte
Immagini (una pagina bianco/nero)	20 Kbyte
Voce (5-10 secondi di parlato)	200 Kbyte

Fig. 1 - Ordine di grandezza tipico per l'occupazione di spazio richiesta da diverse categorie di informazioni [6].

una rappresentazione più strutturata per un testo: riuscendo a definire una struttura non dissimile da quella di una registrazione, si renderebbero applicabili le tecniche di accesso per chiave secondaria.

Il metodo più usuale per associare una struttura ad un testo consiste nell'assegnare una lista di termini, detti *parole chiave,* correlati con il contenuto del testo e tali da distinguerlo dagli altri testi. L'operazione di estrazione delle parole chiave da un testo, denominata indicizzazione (*indexing*), può essere effettuata manualmente dall'utente oppure automaticamente tramite calcolatore; in quest'ultimo caso la procedura fa uso solitamente di diversi archivi di appoggio:

- una lista di esclusione, che riporta le parole comuni che vanno ignorate nella procedura di estrazione delle parole chiave (ad es. "e", "il", "la", etc.);
- una lista di prefissi e suffissi, che aiuta a ridurre una parola nella sua radice (ad. es. pre-, ri-, -mente, -are, -ere, -ire, etc.);
- un dizionario dei sinonimi che serve per assegnare ciascuna radice ad una classe concettuale (ad es. elaboratore = calcolatore).

Nell'indicizzazione si possono usare termini controllati, raccolti in un catalogo o *thesaurus:* con quest'ultimo termine in ambito bibliotecario si indica un insieme di concetti (soggetti o termini) e relazioni fra gli stessi. Il problema di estrarre correttamente le parole chiave più significative di un testo in maniera automatica è comunque lontano dall'essere completamente risolto.

Dopo aver ricavato la lista di parole chiave presenti in un testo, questo può essere assimilato ad una registrazione a lunghezza variabile, fatta da un campo testo e più campi contenenti le parole chiave, il cui numero è molto variabile. Nonostante questa differenza, si possono impiegare anche in questo caso alcuni metodi di accesso multiattributi già usati per le organizzazioni di dati.

Una diversa struttura possibile per i testi è basata sui vettori. Se t è il numero di parole distinte utilizzato in tutti i testi dell'archivio, ogni singolo testo può essere rappresentato da un vettore t-dimensionale, dove una componente non zero significa che una certa parola chiave compare nel testo in questione. Benchè questa rappresentazione sembri maggiormente simile ad una registrazione strutturata, esistono alcune differenze che comportano difficoltà pratiche di realizzazione:

- la dimensione t del vocabolario può facilmente variare con l'inserimento di nuovi testi, mentre gli attributi delle registrazioni di dati sono indipendenti dalla crescita dell'archivio;
- la dimensione di t è molto grande, e raggiunge spesso l'ordine di 10^4-10^5 parole, mentre le normali registrazioni di dati sono costituite da un numero di attributi molto minore, solitamente dell'ordine delle decine o al massimo di qualche centinaio;
- il numero di componenti nulle in questa rappresentazione è assai elevato; nelle registrazioni di dati esse non sono così frequenti.

2.2 Tipi di interrogazioni

I tipi di interrogazioni sui testi possono essere divisi in due categorie. Alla prima appartengono le interrogazioni convenzionali, usualmente applicate per archivi di registrazioni, che prevedono l'impiego di operatori logici AND, OR, NOT, che in questo caso operano sulle parole chiave. Alcuni esempi di interrogazione sono:

- ricerca dei testi contenenti la parola "sistemi";
- ricerca dei testi contenenti la parola "sistemi" oppure la parola "informativi";
- ricerca dei testi contenenti la parola "sistemi" e anche la parola "informativi" (in qualsiasi ordine);
- ricerca dei testi contenenti la parola "sistemi informativi": il testo deve contenere le due parole contigue, nell'ordine specificato, ignorando soltanto spazi, tabulatori, salti a capo.

Alla seconda categoria appartengono interrogazioni che tengono conto della strutturazione del testo in frasi e paragrafi, definita in precedenza. Alcuni esempi di interrogazioni in questo contesto sono:

- ricerca dei testi contenenti la parola "sistemi" e nella stessa frase "informativi";
- ricerca dei testi contenenti "sistemi" seguito da "informativi" nello stesso paragrafo.

La seconda interrogazione è una forma specifica di AND in cui il secondo termine deve comparire soltanto in posizione successiva al primo. L'utilità di prendere in considerazione frasi e paragrafi è evidente:

se si è interessati alla ricerca di testi che trattano l'argomento "sistemi informativi", la prima interrogazione eviterà la selezione di testi nei quali compaiano entrambi i termini, ma in contesti differenti e scorrelati.

Inoltre le interrogazioni testuali considerano solitamente le lettere minuscole e maiuscole alla stessa stregua e permettono di specificare come sequenze di ricerca parole complete o parti di parole. Parti di parole valide sono le seguenti:

- inizio di una parola (ad es. "inf*")
- fine di una parola (ad es. "*atica")
- parte mediana di una parola (ad es. "*mat*)

In questo contesto l'asterisco può corrispondere ad un qualsiasi numero di caratteri, zero incluso. Così tutti e tre gli esempi riportati corrispondono alla parola "informatica", mentre gli ultimi due corrispondono anche a "matematica". È possibile anche far uso del simbolo "?" che sostituisce esattamente un carattere: ad es. "sistem?" può corrispondere a "sistema" o "sistemi", ma non a "sistemistica".

Nonostante questa libertà di interrogazione, è possibile non riuscire a reperire un testo che tratti di un argomento cercato (ad es. "sistemi informativi") per il motivo che questo non compare mai nella forma specificata nel documento. Come risultato, per testi lunghi può capitare di recuperare documenti non rilevanti, non recuperando invece documenti che sono rilevanti, ad esempio a causa della presenza di sinonimi. La rilevanza non è in questo caso un concetto definibile formalmente, ma è un giudizio espresso da chi ha formulato la richiesta.

Il sistema che gestisce i documenti possiede una propria descrizione del contenuto degli stessi: per rispondere ad un'interrogazione deve procedere ad un confronto tra le descrizioni. Poiché la descrizione di un documento non è un dato oggettivo, la descrizione specificata nella richiesta può non essere direttamente confrontabile con nessuna di quelle note al sistema. Obiettivo del sistema è di rendere confrontabili descrizioni diverse di documenti in modo da recuperare tutti i documenti rilevanti con un numero quanto più limitato possibile di documenti irrilevanti.

Si dice *precisione* del sistema di reperimento dell'informazione il rapporto tra i documenti rilevanti ottenuti in risposta ed il numero totale dei documenti recuperati: nel caso di interrogazioni testuali è spesso tollerabile una precisione minore di 1, che significa che nella risposta si prevede un certo numero di documenti non pertinenti (*false drops*), che l'utente provvederà semplicemente a scartare.

Si definisce *richiamo (recall)* di un sistema di reperimento dell'informazione il rapporto tra i documenti rilevanti ottenuti in risposta ed il numero totale di documenti rilevanti presenti nel sistema. Se il rapporto di richiamo è minore di 1 si tollera la possibilità di non reperire tutti i documenti utili (*false dismissals*).

Poichè l'obiettivo di una ricerca testuale è la determinazione di tutti i testi che soddisfano le condizioni di ricerca, lasciando poi all'utente la valutazione dell'utilità per il suo problema dei testi reperiti, si comprende come ci sia spazio per una seconda categoria di interrogazioni, nella quale l'utente richiede tutti i testi che *siano interessanti* per un certo problema, indipendentemente dalla presenza o dall'assenza di una stringa specificata. Si perde in questo caso la separazione netta tra testi che soddisfano o meno un'interrogazione, ma si ordinano invece i testi in base al maggiore o minore interesse nei riguardi di un certo argomento di ricerca, assegnando "pesi" relativi alla significatività di un testo per un certo tipo di ricerca. Questo tipo di interrogazione è oggetto di attenzione in alcuni lavori sul tema del reperimento di informazioni testuali.

2.3 Requisiti operativi

Un requisito importante dei sistemi di reperimento di informazioni testuali è la grande predominanza delle operazioni di ricerca rispetto agli aggiornamenti. Nelle applicazioni di biblioteca non esistono aggiornamenti e le uniche operazioni autorizzate sono l'inserimento e la ricerca (più raramente la cancellazione).

Nelle applicazioni in ambiente d'ufficio, invece, gli aggiornamenti sono permessi, ma rappresentano solitamente una piccola percentuale - dell'ordine del 10% [5] - delle operazioni di ricerca.

Le considerazioni illustrate portano numerose conseguenze per l'organizzazione di grandi archivi testuali:

- il moderato tasso di aggiornamento permette di utilizzare supporti di memorizzazione di massa non riscrivibili;
- si può assumere che l'archivio di testi sia statico, con inserimenti e cancellazioni periodiche che comportano la riorganizzazione delle strutture ausiliarie;
- si può ipotizzare una frequenza di accesso ai testi che è inversamente proporzionale alla loro età.

Non va trascurato, infine, il fatto che gli archivi testuali devono, per loro natura, essere di grandi dimensioni per avere rilevanza nelle applicazioni specifiche. Le operazioni di ricerca su archivi di dimensioni nell'ordine dei Gigabyte sono pertanto limitate da stringenti vincoli di prestazioni.

3. Tecniche di organizzazione per testi

I metodi di accesso ai testi proposti nella letteratura si possono dividere in quattro grandi categorie:

- organizzazioni sequenziali
- organizzazioni a liste invertite
- organizzazioni per chiavi multiple
- organizzazioni basate sul raggruppamento.

Mentre i primi tre metodi sono applicazioni al caso dei testi di organizzazioni studiate per archivi di registrazioni, l'ultima tecnica è stata originariamente introdotta per applicazioni di reperimento di informazioni. Nel seguito verranno presentate le caratteristiche dei metodi più efficienti in ciascuna di queste categorie.

Nella letteratura è stata dedicata in passato anche molta attenzione a soluzioni hardware specializzate per le funzioni di reperimento di informazioni testuali; la rapida crescita delle prestazioni degli elaboratori commerciali rende però difficile ed antieconomico puntare su questa strada, che richiede tempi di sviluppo e costi nettamente superiori e non garantisce l'integrazione con i sistemi informativi aziendali.

3.1 Organizzazioni sequenziali

L'organizzazione più naturale per un archivio di testi è quella sequenziale, che richiede la scansione completa dell'archivio (*full text scanning*) per trovare i documenti che contengono determinate parole chiave. La ricerca viene effettuata confrontando una sequenza di caratteri (*pattern*) con le parole contenute nei testi dell'intero archivio, fino a quando non vengono individuati tutti i documenti che corrispondono a quanto richiesto. Una sequenza di ricerca è valida anche se in essa compaiono i caratteri speciali precedentemente definiti.

Questo metodo non richiede più memoria di quella strettamente necessaria e consente agevoli inserimenti ed aggiornamenti dei testi, giacchè non ha bisogno di strutture ausiliarie da mantenere aggiornate: la ricerca sequenziale all'interno di archivi di testi voluminosi è però molto lenta.

Un algoritmo ovvio per ricercare sequenze di caratteri in un testo si compone di due passi:

- confrontare i caratteri della sequenza di ricerca con i corrispondenti caratteri nel testo;
- se non si verifica la completa coincidenza delle due sequenze, spostare di una posizione verso destra la sequenza di ricerca ed iterare il processo fino a quando si verifica la coincidenza oppure si raggiunge la fine del testo stesso.

Il costo di questo algoritmo è dell'ordine di m^n, se m e n sono rispettivamente le lunghezze della sequenza di ricerca e del testo.

Un algoritmo più veloce è stato proposto da Knuth, Morris e Pratt [9], e richiede un costo dell'ordine di m+n. Questo algoritmo, detto KMP dalle iniziali degli autori, consente di calcolare "in anticipo" di quanti caratteri dev'essere spostata verso destra la sequenza, quando un suo carattere di indice j non coincide con il carattere del testo che è in corrispondenza con esso: lo spostamento verso destra della sequenza di ricerca non avviene più quindi di un solo carattere alla volta, come nel caso precedente, e si avvale in qualche modo dei confronti già effettuati sui caratteri successivi. Il metodo richiede però un tempo di preelaborazione della sequenza di ricerca dell'ordine di m.

Nell'algoritmo KMP la scansione della linea di testo avviene tramite un particolare automa a stati finiti: per ogni nodo, che contiene un carattere della sequenza di ricerca, esistono due informazioni:

- una per l'avanzamento, da seguire nel caso sia stato letto il carattere desiderato (*success link*);
- una per il ritorno indietro, da seguire in caso contrario (*failure link*).

Il criterio di scansione del testo è il seguente:

- il carattere successivo del testo viene letto solo se è eseguito un avanzamento;
- lo stesso carattere del testo viene riconsiderato se è stato eseguito un ritorno indietro, dopo aver spostato la sequenza di ricerca di quanto indicato.

Se, ad esempio, si ricerca la sequenza ABABCB nel testo ACABAABABCBAA, l'algoritmo KMP procede nel modo seguente:

1) *Preelaborazione:* si costruisce l'automa a stati finiti, ad es. tramite una coppia di vettori, il primo, V1, per la sequenza di ricerca, che coincide con la sequenza di avanzamento, e l'altro, V2, per i ritorni indietro. Nel caso dell'esempio si ha

   V1 = (* A B A B C B)

   V2 = (1 0 1 1 2 3 1).

2) *Ricerca:* la scansione del testo avviene come in Fig. 2.

La sequenza di ricerca può essere spostata di più di un carattere alla volta, a seconda di quanto indicato nel vettore V2 in corrispondenza con il carattere della sequenza di ricerca che ha causato il fallimento del confronto.

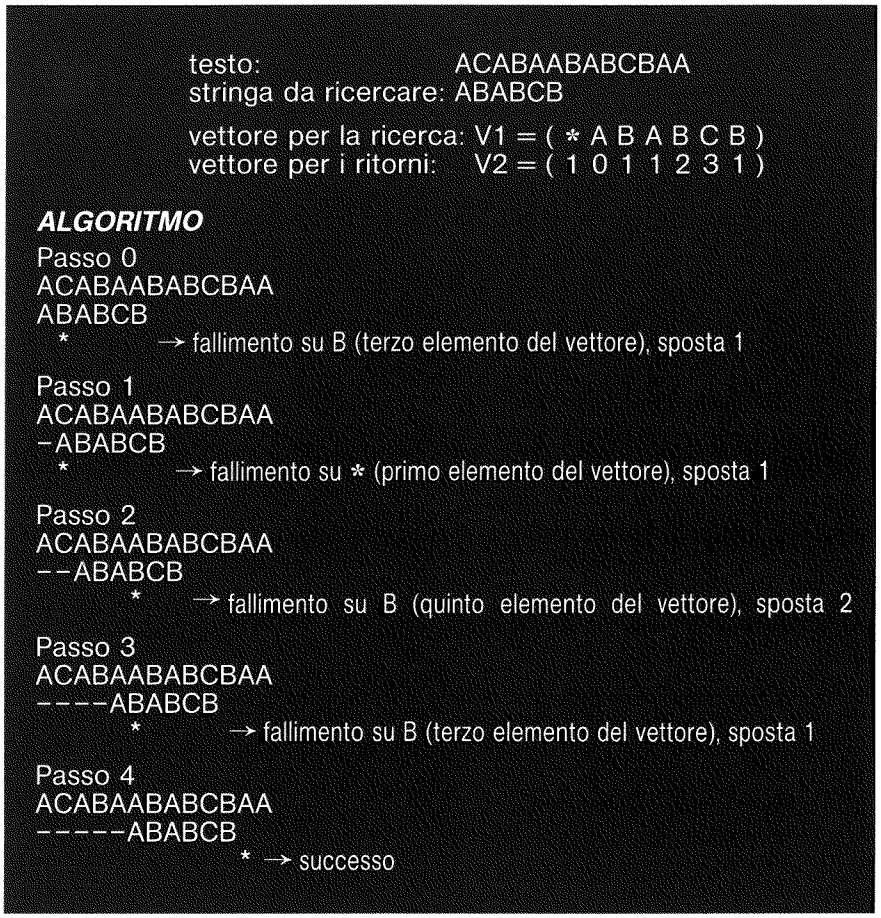
Il più veloce algoritmo conosciuto in questo campo è stato proposto da Boyer e Moore [1]. L'idea che differenzia questo algoritmo dal precedente consiste nell'effettuare i confronti di caratteri da destra verso sinistra: così, se si non si verifica la coincidenza, si può immediatamente spostare verso destra di m caratteri la sequenza cercata. Il numero di confronti è ancora dell'ordine di n+m nel caso peggiore, ma solitamente è molto inferiore, in particolar modo se l'alfabeto contiene molti caratteri. Anche in questo caso è richiesto un tempo di preelaborazione dell'ordine di m.

I vantaggi dell'organizzazione sequenziale si possono riassumere nella semplicità di realizzazione, nell'assenza di strutture dati ausiliarie da aggiornare in caso di modifiche sui testi e nel conseguente risparmio di spazio in memoria di massa. Lo svantaggio è costituito dagli elevati tempi di risposta, che diventano inaccettabili su grandi archivi di testi. Per questa ragione l'organizzazione sequenziale si adopera in pratica soltanto accoppiata ad altri metodi che delimitino la porzione dell'archivio da sottoporre alla scansione completa.

3.2 Organizzazioni a liste invertite

Dopo aver associato ad ogni testo un insieme di parole chiave, un'organizzazione abbastanza ovvia consiste nel costruire una lista invertita su queste chiavi, organizzata ad esempio in ordine alfabetico, che permette di ottenere buoni tempi di risposta in ricerca e permette di trattare con facilità i sinonimi. La presenza dell'indice richiede però una notevole occupazione di memoria (valutata dal 50% al 300% della dimensione dell'archivio di testo

Fig. 2 - Operazioni di ricerca di una stringa in un testo secondo l'algoritmo KMP.



[7], anche se è possibile restringere le parole significative all'interno di un vocabolario prefissato.

L'inserimento di nuovi documenti comporta inoltre un pesante aggiornamento dell'indice e non esistono metodi efficienti per trattare le ricerche con caratteri iniziali non specifici, come ad esempio, "*matica". L'efficacia del reperimento e la facilità di inserimento sono antagonisti nelle tecniche basate sull'inversione: infatti migliori tempi di risposta si possono ottenere con accorgimenti particolari nell'organizzazione dell'indice (ad es. compressione), che rendono però difficili le operazioni di inserimento e cancellazione. Nonostante questi svantaggi, l'inversione costituisce una delle soluzioni adottate di preferenza nelle realizzazioni di sistemi per il trattamento di testi, anche perchè offre parametri di precisione e richiamo uguali a 1.

Basato sulle liste invertite è il più noto esempio di sistema commerciale per la ricerca su testi, lo STAIRS (*Storage and Information Retrieval System*) [8], che inverte tutti i termini che non sono presenti in una lista di parole da escludere.

Oltre ad una semplice gestione di dati strutturati, lo STAIRS permette anche ricerche su un insieme di documenti testuali, nei quali possono essere inserite alcune etichette come "titolo", "testo", "sommario".

La ricerca all'interno delle frasi è resa possibile dalla presenza di un carattere speciale di fine frase.

Le interrogazioni sono espresse in un proprio linguaggio interattivo, che presenta le seguenti caratteristiche:

- operazioni di ricerca nel testo (SEARCH): ammette come operando parole, radici di parole o combinazioni di parole
- operazioni di ricerca nei campi strutturati (SELECT)
- operazioni di visualizzazione dei documenti reperiti con le funzioni precedenti (BROWSE)
- operazioni di ricerca pesata nei testi (RANK): fornisce in risposta documenti "pesati" secondo le specifiche dell'utente.

Lo STAIRS è basato su una gerarchia di archivi a cinque livelli (Fig. 3):

- matrice
- dizionario
- archivio delle frequenze
- indice
- documenti.

Il livello più basso è costituito dall'insieme dei documenti. Il livello successivo, l'indice, contiene un'entrata per ciascun documento, identificata da una chiave univoca che è il numero di documento cui sono associate altre informazioni, come data di creazione, autore, autorizzazioni all'accesso e - naturalmente - un riferimento al documento.

Al livello superiore, l'archivio delle frequenze contiene una registrazione per ogni comparsa di una parola in un documento, che memorizza la parola, il numero di documento, il

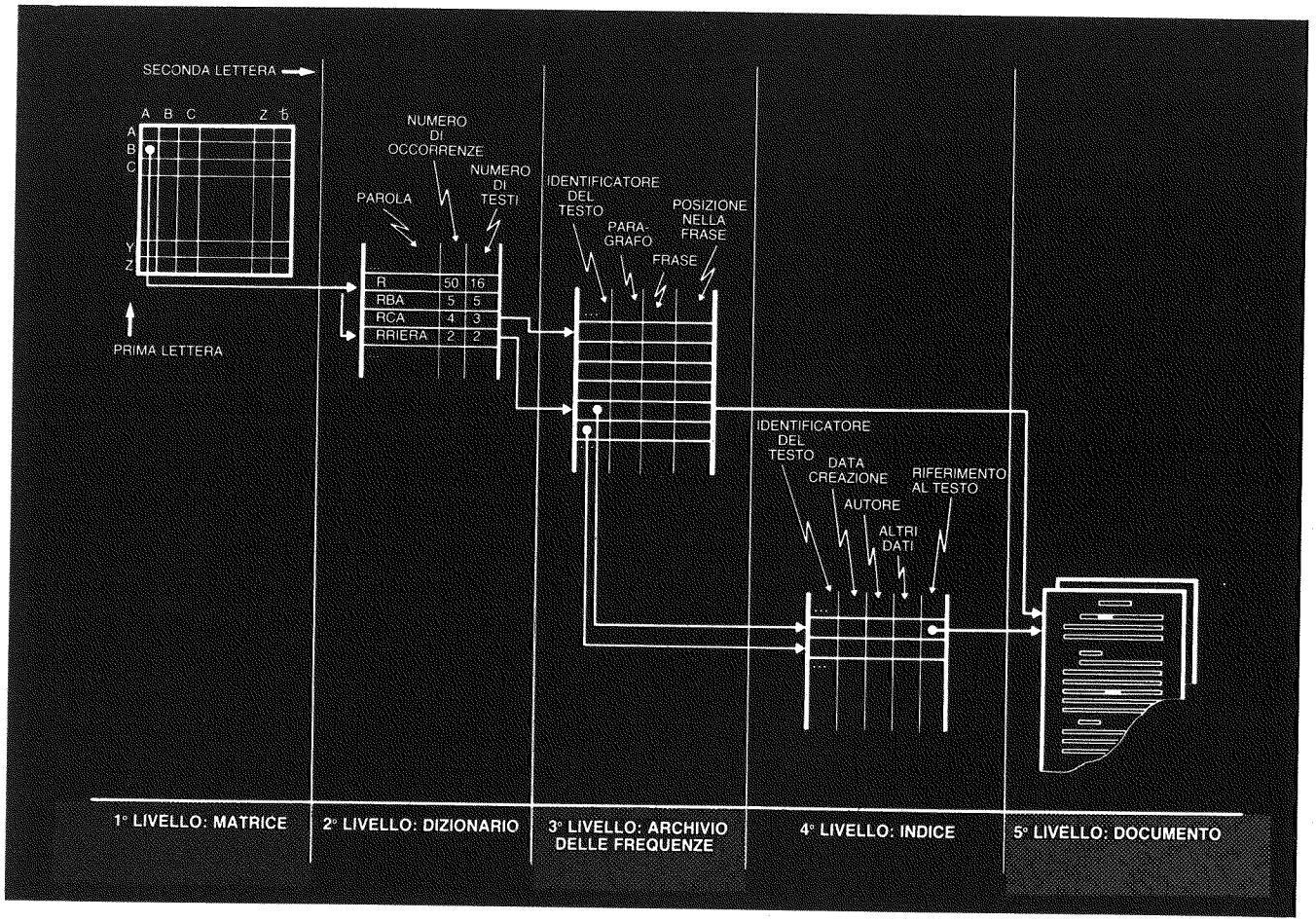


Fig. 3 - Struttura degli archivi dello STAIRS.

codice del paragrafo, il numero di frase e la posizione della parola all'interno della frase. Il riferimento al testo è effettuato tramite il numero di documento.

Poichè una parola ha di solito più entrate nell'archivio delle frequenze, il livello sovrastante, cioè il dizionario, dispone di una registrazione per ogni parola distinta, compresi i termini banali come articoli, congiunzioni, etc. Per risparmiare spazio, i primi due caratteri della parola sono omessi nel dizionario, in quanto ricostruibili dall'elemento di matrice che ad esso fa riferimento.

Per ogni parola memorizzata, il dizionario contiene informazioni sul numero di occorrenze della parola e sul numero di testi diversi nei quali compare.

Al livello più elevato, la matrice, permette un veloce accesso al dizionario: in una tabella di 26x27 entrate sono infatti memorizzate le combinazioni dei primi due caratteri che identificano un puntatore alla prima posizione del dizionario in cui compare una parola che incomincia con quelle due lettere. Il ventisettesimo carattere è lo spazio bianco, in modo da poter indirizzare anche parole formate da un singolo carattere. Si tratta dunque in un *trie* ad un solo livello (cfr. [4]).

La risposta alle interrogazioni in STAIRS avviene facendo uso di questi cinque livelli di indirizzamento. Ad esempio per rispondere all'interrogazione:

SEARCH "Sistemi gestione basidati" NELLA STESSA FRASE,

si intersecano le tre liste per "sistemi", "gestione" e "basidati", per trovare le occorrenze che hanno lo stesso numero di documento, codice di paragrafo, numero della frase, e la parola "sistemi" che precede quella di "gestione", che a sua volta precede "basidati".

Un vantaggio delle strutture di dati dello STAIRS consiste nella possibilità di ottenere importanti informazioni senza accedere ai documenti, spesso mantenendosi ai livelli più elevati della gerarchia: esaminando i primi due livelli è possibile conoscere, ad esempio, se una parola è presente nel dizionario; dai primi tre livelli si può trovare quante volte ed in quanti documenti compare una parola cercata. La struttura si presta bene, inoltre, all'integrazione con un *thesaurus*.

Disegnato principalmente per fornire buone prestazioni in ricerca, in particolare nel caso di ricerche su una singola parola o comunque poco complesse, lo STAIRS necessita di una rigenerazione degli indici in

caso di aggiunta o aggiornamento di documenti e pertanto non si adatta con facilità ad un ambiente dinamico.

3.3 Organizzazioni multichiavi

Oltre all'inversione, per accedere ai testi sono state studiate organizzazioni più sofisticate per chiavi multiple (*multiattribute retrieval*). Tra queste si distinguono per la loro utilità le tecniche che fanno uso dell'archivio delle firme (*signature file*). In questo caso si gestisce un archivio di documenti in formato originale ed un archivio di abbreviazioni della parte testuale, sul quale si inizia la ricerca per scartare i documenti che certamente non soddisfano la condizione; la ricerca procede poi sull'archivio dei testi completi. Il metodo presenta richiamo uguale a 1 e precisione minore di 1.

Metodi di estrazione delle firme

Il principio di estrazione delle firme può essere applicato tanto ad una registrazione strutturata, quanto ad un documento (Fig. 4). Al fine di estrarne la firma, ciascuna registrazione dell'archivio viene elaborata separatamente. Ciascun attributo della registrazione (parola, nel caso di un testo) viene trasformato - ad esempio usando una funzione *hash* - in una sequenza di bit, che ne costituisce la firma (*word signature*). Le firme delle parole così ottenute vengono combinate (ad es. concatenate) producendo così la firma della registrazione, che costituisce una sequenza di bit in genere molto più corta della registrazione stessa.

Una volta prodotta la firma di una registrazione, questa può essere utilizzata come indirizzo di memorizzazione, secondo tecniche di *hash multiattributo,* oppure memorizzata sequenzialmente in un archivio, detto *archivio delle firme (signature file).*

- Metodo "firma delle parole"

Per quanto riguarda in particolare l'estrazione delle firme da archivi testuali, l'algoritmo di generazione viene solitamente applicato alle sole parole significative contenute nel testo, cioè con l'esclusione di articoli, congiunzioni, etc. Naturalmente il procedimento di estrazione non è invertibile, cioè non è possibile ricostruire il testo partendo dalla firma, che occupa mediamente il 10% dello spazio richiesto dal documento originale [3]: i testi completi pertanto devono essere comunque memorizzati.

Un primo metodo di estrazione delle firme dai testi è detto *firma delle parole* (*Word Signature* o WS) [14], [10]. Si tratta di un'applicazione che permette di passare dal documento alla sua abbreviazione, preservando l'ordine e l'eventuale duplicazione delle parole.

In questo caso, le firme sono sequenze binarie di lunghezza fissa f (che costituisce uno dei parametri del disegno, ad es. 16 bit), ottenute applicando un'opportuna funzione *hash* a tutte le parole significative del testo, mentre la firma del testo si ottiene semplicemente concatenando tutte le firme delle parole.

Un esempio di applicazione del metodo *firma delle parole* con f=4 per estrarre la firma dal testo "Metodi di estrazione delle firme dei testi" è riportato in Fig. 5:

- il primo passo consiste nell'identificare le parole significative, eliminando articoli e preposizioni articolate, ottenendo così la sequenza "Metodi estrazione firme testi";

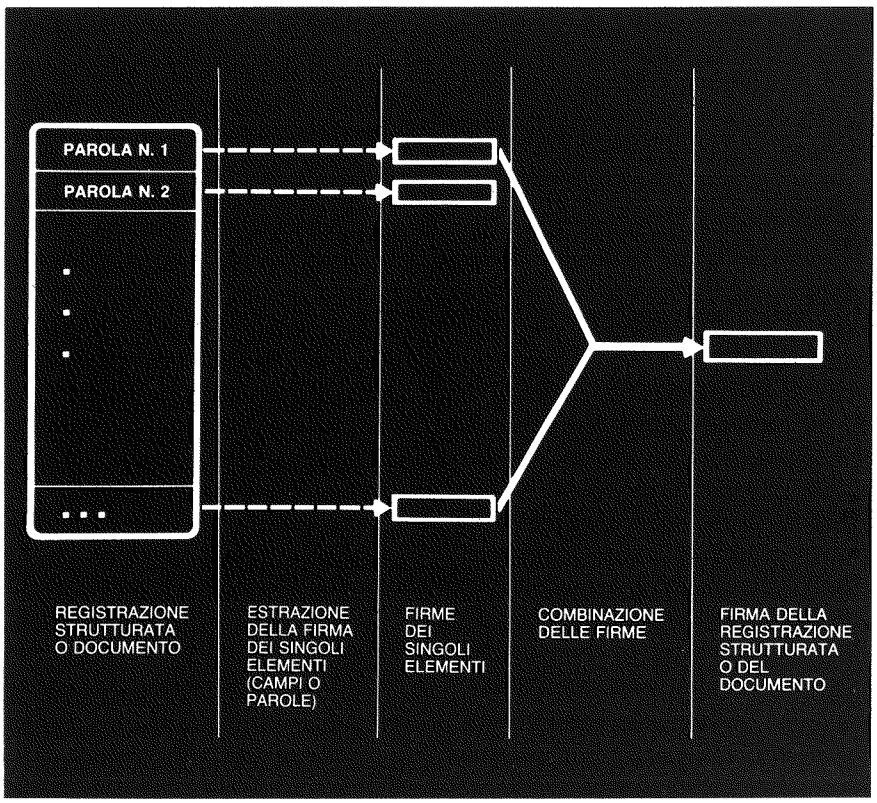


Fig. 4 - Estrazione della "firma" di una registrazione strutturata o di un documento.

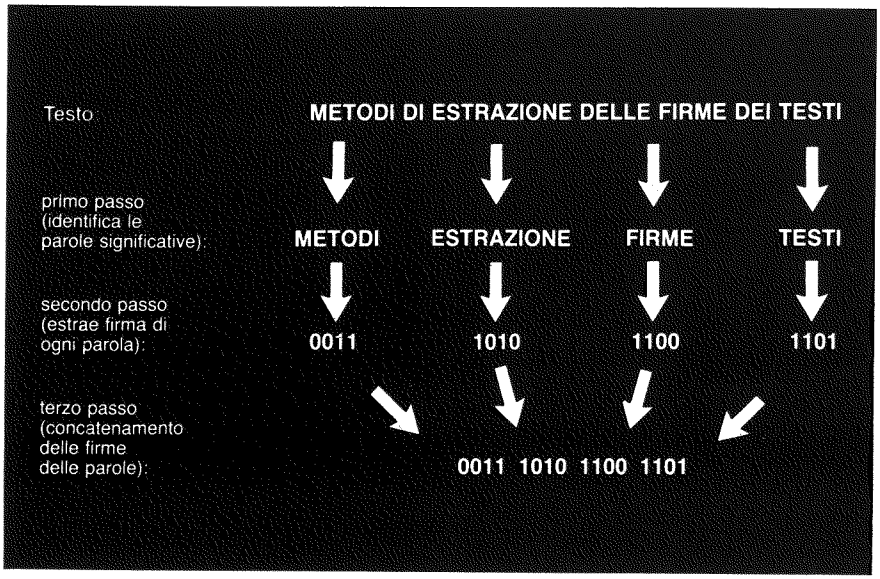


Fig. 5 - Estrazione della firma di un testo con il metodo "firma delle parole".

Parola	Firma			
METODI	001	000	110	010
ESTRAZIONE	000	010	101	001
FIRME	001	010	001	010
TESTI	000	000	111	001
Firma del blocco	001	010	111	011

Fig. 6 - Estrazione della firma di un blocco di testo secondo il metodo della *codifica sovraimposta:* si suppone che ogni blocco logico consista di D=4 parole, che la lunghezza della firma di ogni parola (e quindi di ogni blocco) sia f=12 e che ogni parola ponga a uno m=4 bit.

- la firma di ciascuna parola si ottiene a questo punto applicando un'opportuna funzione *hash*, non specificata in questa sede;
- la firma complessiva del testo si ottiene concatenando le quattro firme delle singole parole.

• Metodo "codifica sovraimposta"

Un secondo metodo di estrazione delle firme è detto *codifica sovraimposta* (*Superimposed Coding* o SC) [2]. Si tratta di un'applicazione che permette di passare dal documento alla sua abbreviazione, perdendo l'ordine delle parole ed ignorando la loro duplicazione all'interno di uno stesso blocco.

In questo caso, infatti, ciascun documento viene diviso in *blocchi logici;* un blocco viene definito come una porzione di testo che contiene un numero prefissato D di parole significative distinte.

Si estrae poi la firma di ogni parola significativa, applicando una funzione che fornisce come risultato m numeri nell'intervallo (1,f); le posizioni corrispondenti del vettore vengono poste a 1. A questo punto si esegue l'OR logico delle firme di tutte le parole all'interno di un blocco (Fig. 6), ottenendo dunque per esso una sequenza di f bit. La firma complessiva del testo si ottiene concatenando infine le firme di tutti i blocchi.

Questo metodo, le cui prestazioni sono caratterizzate da un'accurata scelta dei parametri D, f e m, presenta il vantaggio di permettere di eliminare automaticamente la duplicazione di parole all'interno di un blocco di testo, che con il metodo WS può essere rimossa soltanto effettuando un ordinamento.

Per la scelta dei parametri di progetto, f si sceglie solitamente in base alla lunghezza della parola del calcolatore (16, 32,...). La dimensione del blocco dipende dall'unità di trasferimento e da questa si stima D, numero di parole significative distinte in un blocco.

La *codifica sovraimposta* è in grado di gestire sinonimi, se questi comportano la stessa firma; ciò richiede però di consultare un dizionario, la cui costruzione e manutenzione può risultare onerosa.

Criteri di ricerca

Le firme dei documenti - estratte secondo uno dei metodi esposti e memorizzate nell'archivio delle firme - vengono usate per stabilire in modo veloce se i testi da cui sono estratte possano contenere o meno alcune parole chiave cercate.

La ricerca di una parola in un testo avviene mediante scansione sequenziale nella sua firma: se questa

viene riconosciuta, la firma soddisfa il criterio di interrogazione. Un buon metodo di estrazione deve essere progettato in modo da garantire che se un termine è presente nel testo, la sua presenza sia rilevata dall'esame della firma (richiamo uguale a 1). Non vale usualmente l'opposto, cioè non è detto che le parole siano effettivamente presenti nel testo anche se le firme lo fanno ritenere, a causa delle collisioni tipiche dei metodi *hash* (precisione P minore di 1).

La ricerca nell'archivio delle firme porta dunque a dichiarare soddisfacente alle specifiche un sovrainsieme dei testi richiesti: i testi superflui possono venir eliminati eseguendo successivamente una scansione completa nel testo oppure lasciando all'utente il compito di scartare quelli non validi.

Dovendo, ad esempio, selezionare tutti i testi che contengono un sottoinsieme specificato di parole chiave, si attua per ciascuno di essi una scansione della firma secondo un algoritmo di ricerca delle parole chiave coerente con il processo di estrazione e si stabilisce se il testo può contenere i termini voluti.

Nell'esempio effettuato a proposito della tecnica WS, la ricerca del termine "Metodi" si esegue ricercando la firma della parola, nell'esempio, 0000, all'interno della firma complessiva del documento: se la firma cercata compare nella firma del testo, questo viene dichiarato soddisfacente alla ricerca.

Con la tecnica SC, dopo avere ottenuto la firma della parola cercata, si esamina sequenzialmente la firma del documento; si dichiara che la firma soddisfa la ricerca (e si arresta la scansione) se la firma di uno dei blocchi del testo contiene gli stessi m bit a 1 rispetto alla firma della parola cercata.

Il principale svantaggio dell'archivio delle firme sta comunque nel peggioramento dei tempi di risposta al crescere delle dimensioni dell'archivio. I vantaggi, al contrario, sono la semplicità di realizzazione e la possibilità di soddisfare interrogazioni su prefissi di parole. I metodi di questo tipo richiedono molte elaborazioni in memoria centrale, e sono dunque *CPU-bound;* in [12] si mostra che il metodo più veloce nella scansione della firma è la codifica sovraimposta.

3.4 Organizzazioni basate sul raggruppamento

L'idea di base di questo metodo consiste nel raggruppare insieme i documenti in base a determinati criteri di similitudine, formando così gruppi (*cluster*); ciò si giustifica ipotizzando che testi dello stesso gruppo siano utili in risposta alle stesse richieste, e che il loro raggruppamento acceleri il processo di ricerca [15].

Generazione dei gruppi

La costruzione dei gruppi procede nel modo seguente: ciascun documento viene elaborato trascurando le parole comuni, estraendo i termini significativi e riducendoli alla loro radice.

Ciascuna radice viene ricondotta ad una classe concettuale, definita come insieme di termini sinonimi: orientativamente si può pensare a circa 10.000 classi per un archivio di 100.000 parole.

A questo punto un testo è rappresentato da un vettore nello spazio k-dimensionale delle classi concettuali ammissibili; ciascuna componente del vettore indica con 0 l'assenza della parola chiave associata a quella coordinata e con 1 la sua presenza. In alternativa alla rappresentazione con un vettore di cifre binarie, è stato proposto di indicare la presenza di una parola chiave con un "peso" riferito all'importanza del termine nel contesto del documento. Come "peso" si può usare la frequenza del termine nel documento o una più complessa funzione costruita comunque a partire dalla frequenza.

Stabiliti gli eventuali pesi, è necessario assegnare i documenti ai gruppi, in modo da costruire i raggruppamenti dei documenti secondo un criterio che soddisfi i seguenti requisiti:

- stabilità rispetto alla crescita del volume dei documenti
- scarsa influenza di piccoli errori nella definizione delle parole chiave dei documenti
- indipendenza dall'ordinamento iniziale dei documenti
- efficienza, in termini di tempo richiesto per la costruzione dei raggruppamenti.

I metodi proposti nella letteratura privilegiano ora l'uno ora l'altro di questi requisiti: ad es. i metodi basati sulla *matrice di somiglianza* [13] sono molto validi dal punto di vista teorico, ma richiedono un tempo di costruzione dell'ordine di n^2, dove n è il numero dei documenti presenti.

Questi metodi prevedono la scelta di una funzione di somiglianza, che misura la correlazione tra due documenti; ciò può essere fatto in maniera automatica, ed esempio verificando quante parole chiave hanno in comune. Il partizionamento dei documenti in gruppi dipende dalla funzione scelta e dalla "soglia" di somiglianza fissata: appartengono allo stesso gruppo tutti i documenti con coefficiente di similitudine superiore ad una soglia prefissata.

Al contrario, i metodi iterativi hanno caratteristiche più euristiche e dipendono in generale dall'ordine in cui i documenti sono presenti, ma richiedono un tempo di costruzione minore, dell'ordine di n log n.

L'approccio generale consiste nel determinare un partizionamento iniziale e successivamente riassegnare in maniera iterativa alcuni documenti ad altri gruppi. Una possibilità da non trascurare, in ambiente d'ufficio, consiste nell'utilizzare le valutazioni degli utenti del sistema per modificare la composizione dei gruppi, spostando documenti in funzione dell'utilità che i raggruppamenti presenti hanno dimostrato nelle ricerche effettuate.

Tra i vantaggi di questo metodo si segnala l'ordinamento per similarità decrescente dei documenti nella risposta e la possibilità di retroazione sulla rilevanza: l'utente può segnalare i documenti rilevanti fra quelli trovati ed il sistema riformula il vet-

tore della richiesta sommando i pesi dei vettori dei documenti rilevanti e sottraendo i pesi dei vettori dei documenti non rilevanti.

Il principale svantaggio del metodo consiste nel fatto che non si presta per archivi con alta volatilità, in quanto l'inserimento di un documento che garantisca un buon criterio di raggruppamento è piuttosto costosa.

Ricerca nei gruppi

Al termine dei processi descritti, i documenti si presentano raggruppati insieme in base alla loro somiglianza. La ricerca nei gruppi è molto più semplice della loro generazione.

La richiesta viene rappresentata da un vettore N-dimensionale che viene confrontato con i centroidi dei gruppi, a partire da quelli più vicini, secondo la metrica assegnata. In questo modo è possibile determinare i gruppi sui quali è più promettente continuare la ricerca dei testi, fino a fornire all'utente una risposta che specifica una lista di documenti in ordine di importanza decrescente, fino ad un valore di soglia prefissato.

Riferimenti bibliografici

[1] BOYER R.S. e MOORE J.S., *A fast string searching algorithm,* Commun. ACM, *20(10)*, 1977.

[2] CHRISTODOULAKIS S. e FALOUTSOS C., *Design considerations for a message file server,* IEEE Trans. Software Eng., *1/(2)*, 1984.

[3] FALOUTSOS C., *Access method for text,* ComputinS Surveys, *17(1)*, 1985.

[4] FELICIAN L., *Una struttura dati che realizza l'accesso per caratteri: il trie,* Note di Software, 33, 1986.

[5] FELICIAN L., *Image Base Management Systems: a promising tool in the large office system environment,* Database, *19(1)*, 1988.

[6] FELICIAN L. e SCARAMUZZA R., *A multimedia database prototype in the personal computer environment,* Proc. 1st EURINFO Conference, 1988.

[7] HASKIN R.L., *Special purpose processors for text retrieval,* Database Eng. *4(1)*, 1981.

[8] IBM, *STAIRS/VS Reference manual,* 1979.

[9] KNUTH D.E., MORRIS J.H. e PRATT V.R., *Fast Pattern matching in strings,* SIAM J. Comput, *6(2)*, 1977.

[10] LARSON P.A., *A method for speeding up text retrieval,* Proc. ACM SIGMOD Conf. 1983.

[11] McILORY M.D., *Development of a spelling list,* IEEE Trans. Commun., *30(1)*, 1982.

[12] RABITTI F. e ZIZKA J., *Evaluation of access methods to text documents in office systems,* Proc. ACM-BCS Symp. Res. Devel. Info. Retrieval, 1984.

[13] SALTON G. e McGILL M.J., *Introduction to modern Information Retrieval,* McGraw Hill, New York, 1983.

[14] TSICHRITZIS D., *Message Files, ACD,* TOIS, *1(1)*, 1983.

[15] VAN RIJSBERGEN C.J., *Information Retrieval,* 2nd edition, Buttherworths, London, 1979.

Computer e inquinamento elettromagnetico

MAURIZIO MARTINENGO
Bull Italia
Formazione Specialistica Interna
Milano

1. Introduzione

La Compatibilità Elettromagnetica (Electro Magnetic Compatibility, E.M.C. nel seguito) è uno degli aspetti più importanti dell'elettronica moderna, sia dal punto di vista progettuale/commerciale che dal punto di vista installazione e manutenzione di impianti comprendenti apparecchiature elettroniche/elettriche di qualsiasi tipo. Scopo di queste note è di focalizzare le problematiche fisiche generali E.M.C. e, segnatamente, gli impatti tecnici relativi alle macchine ed agli impianti per l'Informatica. Definiremo il concetto di Compatibilità Elettromagnetica partendo da quello, opposto, di Interferenza Elettromagnetica (Electro Magnetic Interference, E.M.I. nel seguito). Si può definire "Interferenza Elettromagnetica" qualsiasi segnale elettrico che, durante il funzionamento di un'apparecchiatura elettronica, si sovrapponga ai suoi segnali digitali o analogici variandone i parametri elettrici ed il contenuto informativo (Fig. 1 e 2). Gli oggetti che sono fonte di E.M.I. (sources) e gli oggetti colpiti da interferenza (victims) sono parte del cosidetto "ambiente elettromagnetico", inteso come la regione di spazio (cabinet di apparecchiatura, sala calcolo, edificio, ecc...) in cui le Interferenze sono riscontrabili in base ai loro effetti. La corretta definizione dell'"ambiente" che fa da contorno a un problema E.M.I. è una delle premesse importanti alla sua soluzione pur costituendo una fase assai impegnativa di ricerca, specie nel caso di interferenze in impianti.

Si può perciò definire la "Compatibilità Elettromagnetica" come una disciplina che analizza le Interferenze citate e fornisce gli strumenti teorici e pratici per ridurre un "Ambiente Interferente" ad un "Ambiente Elettromagneticamente Compatibile", agendo sulle condizioni al contorno e sulle apparecchiature in esso contenute. Ciò significa, in ultima analisi, ridurre l'emissione di disturbi e/o i valori di sensibilità degli apparati entro i necessari limiti.

È ora opportuno esaminare più nel dettaglio la duplice natura dell'Interferenza Elettromagnetica.

2. Interferenze irradiate e condotte

È bene premettere che il sempre più frequente manifestarsi di Ambienti Interferenti è, in alcuni casi, in rapporto con l'evoluzione della microelettronica, giunta attualmente ad un elevatissimo grado di densità circuitale (10^6MOS/chip). La conseguente concentrazione di componenti e collegamenti in spazi molto ristretti è spesso fattore scatenante di problemi E.M.I.

2.1 Interferenze irradiate

Il concetto di interferenza irradiata è legato al concetto stesso di Campo Elettromagnetico (campo e.m.). È noto dalla fisica generale che una qualsiasi carica elettrica in moto accelerato (ad esempio uno o più elettroni in moto in un conduttore con velocità variabile) irradia energia tramite una grandezza fisica di natura ondulatoria detta Campo Elettromagnetico, costituita da un campo elettrico E e da un campo magnetico B variabili nel tempo con linee di forza tra loro interconnesse (Fig. 3).

La corrente variabile costituisce la sorgente del campo e.m., il quale è in grado di propagarsi autonomamente lontano dalla sorgente nello spazio circostante (Fig. 4). Il campo e.m. trasporta energia e, su eventuali circuiti elettrici colpiti detti "antenne riceventi", è in grado di replicare in forma e frequenza la corrente originaria. L'intensità di tale effetto dipende principalmente dalla geometria (lunghezza) dell'antenna verso la lunghezza d'onda del campo, dall'attenuazione che il campo e.m. subisce durante la sua propagazione nonchè dalla sua intensità alla sorgente (potenza di emissione).

Si può ora meglio definire l'Interferenza Irradiata come un campo e.m. (onda e.m.) che colpisce un circuito elettrico non destinato a riceverla, disturbandone con i suoi effetti il funzionamento.

Questa interferenza si manifesta sia in ambiente aperto che in ambiente chiuso.

2.2 Interferenze condotte

Le interferenze condotte si propagano sui conduttori delle tensioni di rete e sui conduttori di massa ad essi associati, e sono perciò tipiche dell'ambiente chiuso.

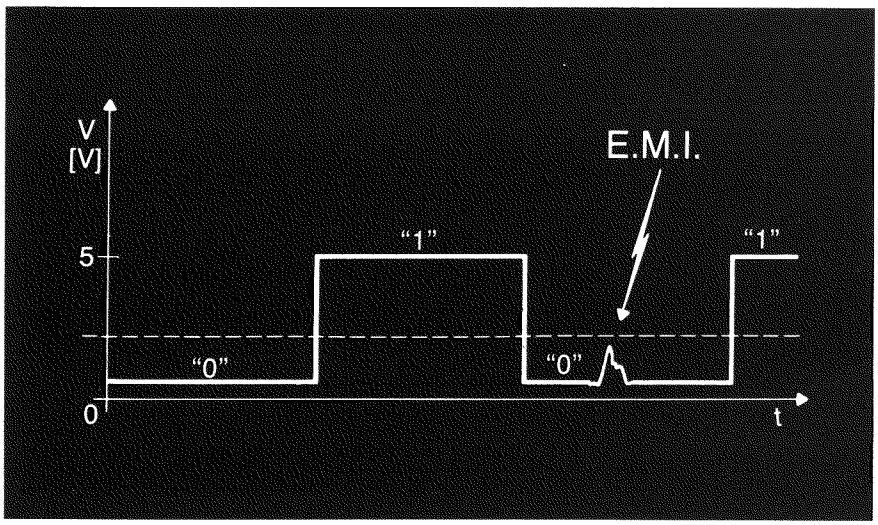


Fig. 1 - Assenza di disturbi, o disturbi "compatibili".

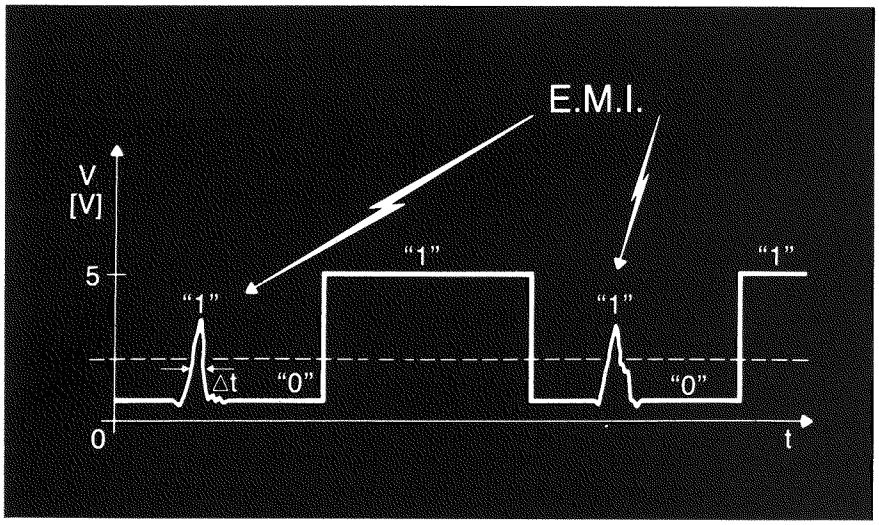


Fig. 2 - Disturbi su livelli di "0" logico che superano la soglia di "noise immunity" della famiglia logica, con conseguenti errori. La commutazione 01 spuria necessita anche di una certa durata ΔT dell'impulso (dipendente dalla logica) per generare interferenza.

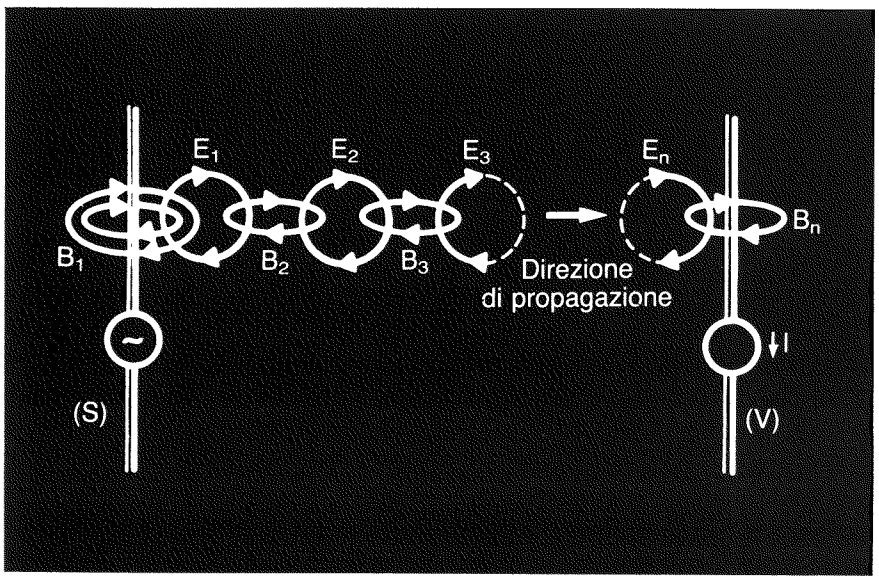


Fig. 3 - Struttura di un'onda elettromagnetica e sua propagazione tra circuito sorgente (S) e vittima (V).

Fig. 4 - Propagazione di un'onda elettromagnetica generata da una corrente sinusoidale. Questo tipo di onda e.m. si definisce "piana" perchè la componente elettrica e quella magnetica variano in ampiezza su direzioni perpendicolari a quella di propagazione.

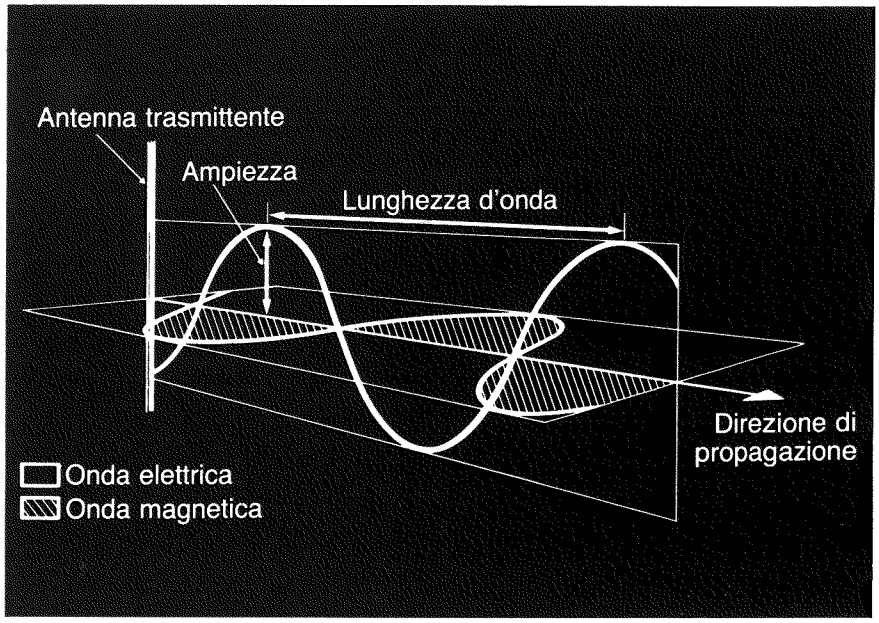
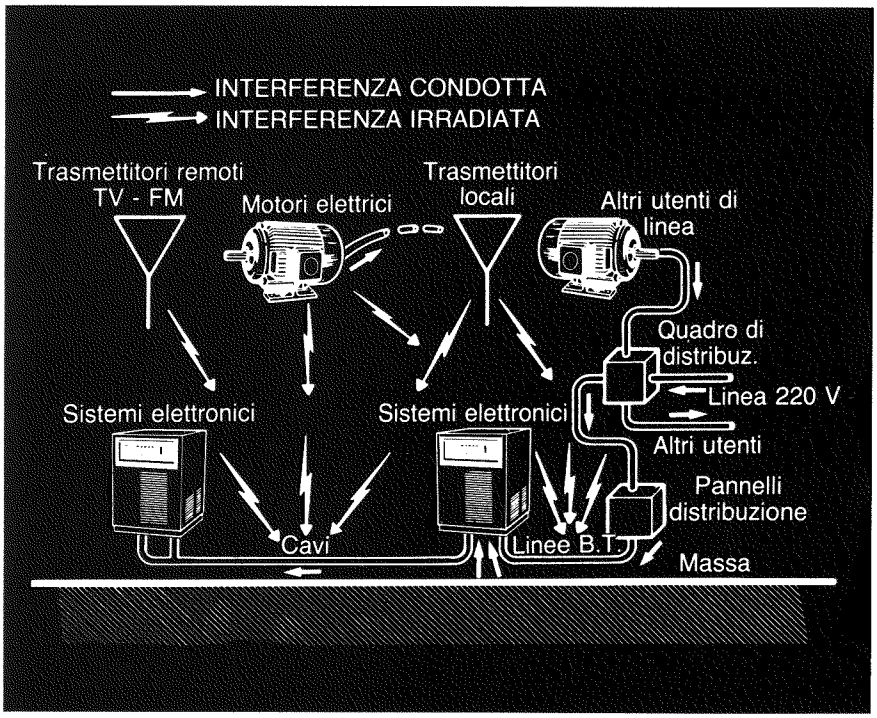


Fig. 5 - Disturbi condotti ed irradiati.



La fig. 5 sintetizza in modo immediato la situazione di alcuni apparati sorgente (sources) e di uno suscettibile (victim), a fronte di interferenze che si propagano nei due modi base considerati, cioè per via irradiata e per via condotta.

2.3 Elementi interferenti in sistemi EDP

L'elenco seguente riporta, a scopo esemplificativo, alcuni sottosistemi e/o dispositivi tipici delle macchine per elaborazione dati (intese come unità centrali, periferiche e componenti di esse) interessati al ruolo di sources o victim tipici a livello di E.M.I.

Sources: Clocks
Alimentatori switching
Line drivers
Amplificatori di scansione
Pennelli elettronici
Unità video
Testine di stampa

Victims: Bus digitali ad alta impedenza
Amplificatori per dischi (e nastri)
Amplificatori video
Connettori I/O dati
Convertitori Analogico/Digitale

3. Tassonomia dell'Interferenza Elettromagnetica

Un corretto approccio alla Compatibilità Elettromagnetica prevede come base di partenza su cui lavorare dinnanzi ad un problema E.M.I. la classificazione dell'interferenza presente, facendo riferimento alle "interferenze fondamentali" sotto riportate"

È però doveroso osservare che molto spesso un problema reale non si presenta come una interferenza "pura", bensì come un insieme di due o più interferenze fondamentali.

Le metodologie di ricerca del tipo di interferenza esulano dallo scopo di questo articolo; si farà comunque cenno, laddove possibile, alle necessarie considerazioni.

3.1 Interferenza per massa comune (C.G.C.)

La sigla che la contraddistingue è l'acronimo della dizione inglese "Common Ground Coupling", e sta ad indicare interferenza per impedenza di massa comune.

La Fig. 6a è esemplificativa del caso: in essa, i due apparati denominati Box 1 e Box 2 hanno i conduttori di ground connessi tra loro tramite il conduttore di massa comune indicato. Circa la natura dei Box, quelli in figura rappresentano intuitivamente due apparati di un sistema EDP (ad esempio Box 1 è host e Box 2 un terminale); nulla cambia nella natura del problema se si pensano i Box come due dispositivi elettronici interconnessi su una piastra, all'interno di una unità qualsiasi.

L'interferenza vera e propria consiste nella possibile generazione di una tensione ai capi del tratto AH, qualora una corrente Ig (di natura ad esempio impulsiva - Fig. 6b -) percorra tale tratto di conduttore. Appare chiaro come il valore di riferimento di massa dei circuiti di Box 2 (punto H) si "innalzi" del valore:

$$Vi = Z \cdot Ig$$

rispetto a quelli di Box 1 (punto A).

Il valore di Vi può causare E.M.I. (ad esempio errori nel riconoscimento di livelli elettrici di "0" e "1") nei circuiti del Box 2, e quindi nel sistema. Questo accade quando si eccede la soglia massima di Immunità (sul livello di zero logico) prevista dalla tecnologia dei componenti elettronici impiegati (Fig. 2). Si tratta di una interferenza tipica di "ambiente chiuso".

3.2 Interferenza di modo comune (C.M.C.)

La sigla che la contraddistingue è l'acronimo della dizione inglese "Common Mode Coupling", e sta ad indicare interferenza per accoppiamento di modo comune su loop di massa (cioè sul circuito chiuso A-H-K-Z di Fig. 7 e 6).

L'esame della Fig. 7 rivela inoltre come questa interferenza si manifesti su una architettura Hardware identica a quella esaminata per il C.G.C., in presenza però di un'onda e.m. incidente sul loop di massa dall'esterno.

Si impongono a questo punto due considerazioni: la necessità della presenza di un campo irradiato da altro dispositivo rende questa interferenza possibile in "ambiente aperto"; è inoltre ancor più evidente la pericolosità di connessioni di massa come quelle delle figure 6a e 7, che sommano a possibili problemi di C.G.C. anche problemi di cattura di campi interferenti emessi da altri dispositivi (C.M.C. appunto).

L'antenna, per così dire, è costituita del loop retinato in Fig. 7 e le correnti interferenti che circolano nei conduttori del loop (I_1 e I_2) sono equiverse (modo comune) nel cavo di interconnesione. Ciò consente particolari approcci risolutivi.

Tra gli esempi riportabili è possibile citare il caso (in ambiente aperto) di C.M.C tra il sistema radio cercapersone di un'azienda (source) e la rete di ground della sua sala calcolo (victim) ed il caso di C.M.C. (in ambiente chiuso) tra un clock di sistema (source) ed un loop di massa su una piastra vicina (victim) all'interno di un cabinet di apparecchiatura.

3.3 Interferenza di modo differenziale (D.M.C.)

La sigla che la contraddistingue è l'acronimo della direzione inglese "Differential Mode Coupling", e sta ad indicare interferenza per accoppiamento di modo differenziale su cavo di interconnesione.

L'elemento vittima è, in questo caso, il cavo che collega i box di Fig. 8 (area retinata). Le correnti I_1 e I_2 non sono equiverse ed i metodi di soppressione sono diversi dal caso C.M.C.

È molto importante, comunque, la prevenzione di questa interferenza mediante una accurata scelta del tipo e delle connessioni di ground agli estremi del cavo, note le caratteristiche dell'ambiente elettromagnetico.

3.4 Interferenza per cross-talk (C.C.C.)

La sigla che la contraddistingue è l'acronimo della direzione inglese "Cable to Cable Coupling", e sta ad indicare interferenza per cross-talk (capacitivo o induttivo) tra cavi vicini.

È la più classica interferenza tra conduttori con geometrie simili e distanze ridotte, (Fig. 9) e consiste nel passaggio di energia tra il cavo source ed il victim per via campo e.m.,; tuttavia, le particolari geometrie consentono di semplificarne l'approccio interpretando il fenomeno come un puro accoppiamento capacitivo o induttivo, mediante precise procedure di calcolo.

Tali procedure consentono di pilotare in modo opportuno la riduzione del disturbo.

4. Standards E.M.C. e vincoli commerciali

L'obiettivo tradizionale di una normativa circa questi fenomeni è sempre stato quello di limitare i disturbi emessi dagli apparati elettrici ed elettronici, allo scopo di eliminare le interferenze con le trasmissioni Radio-TV.

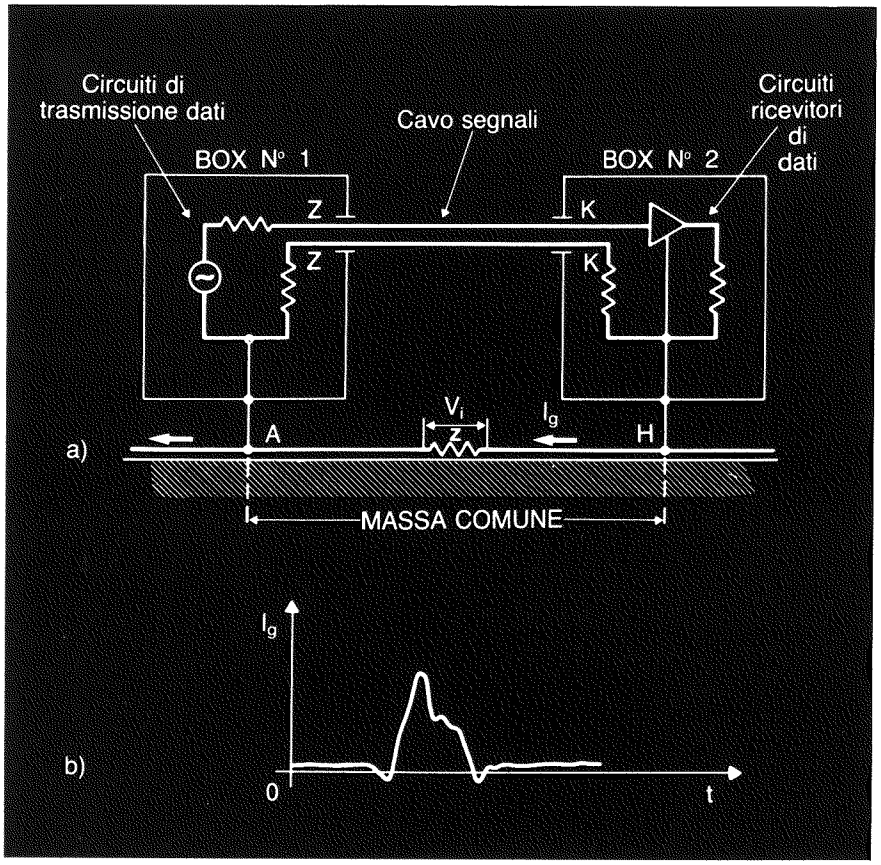


Fig. 6 - a) Accoppiamento per massa comune (C.G.C.)
b) Corrente di massa di tipo impulsivo (transitorio). In questa figura e nelle seguenti non sono riportati i cavi di alimentazione da rete.

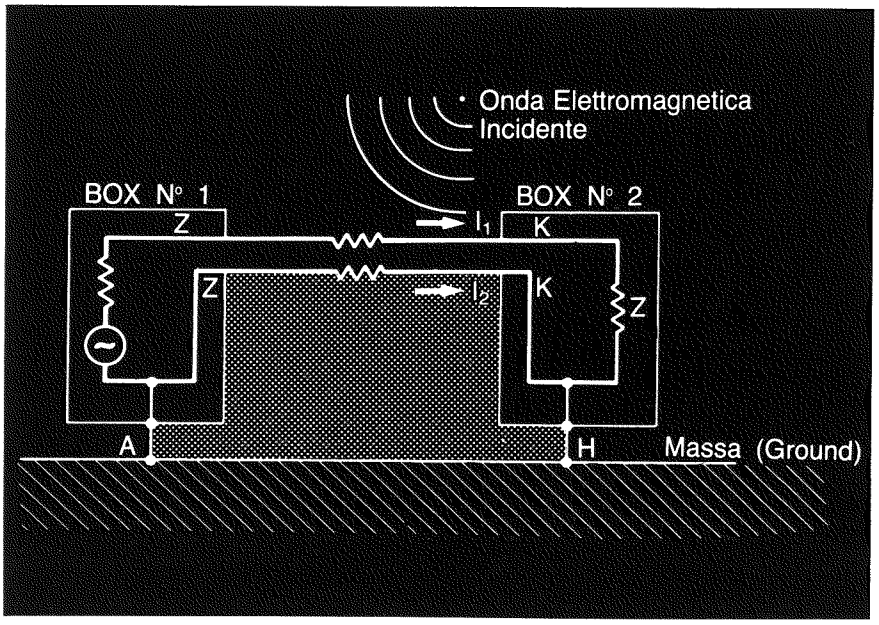


Fig. 7 - Accoppiamento su loop (C.M.C.).

A ciò si sono aggiunti i nuovi orientamenti degli Enti Professionali del settore in merito alla protezione di differenti servizi (trasmissione dati, applicazioni EDP, ecc...), alla garanzia della compatibilità di più sistemi elettrici ed elettronici coesistenti nello stesso ambiente (casa, ufficio, ecc...) e, non ultima, alla definizione di Standard relativi ai nuovi apparati multifunzione (es. TV dotati di stampanti).

La situazione attuale, alquanto diversificata a livello internazionale, è brevemente accennata nei paragrafi seguenti.

È quasi superfluo osservare l'estrema importanza di tali Standards per le aziende che, come la Bull Italia, progettano e costruiscono macchine per l'EDP destinate ad essere commercializzate sui mercati nazionali ed esteri.

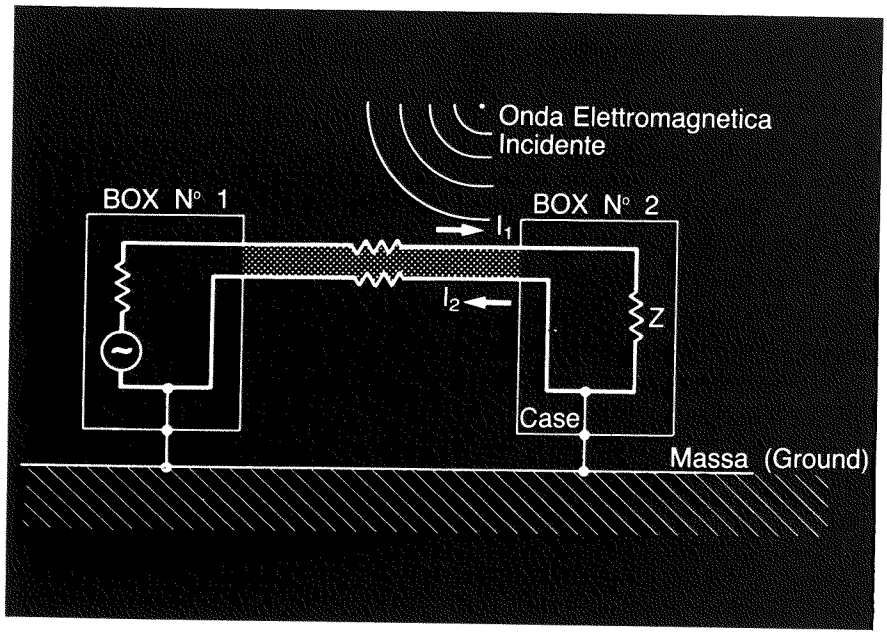


Fig. 8 - Accoppiamento su cavo (D.M.C.).

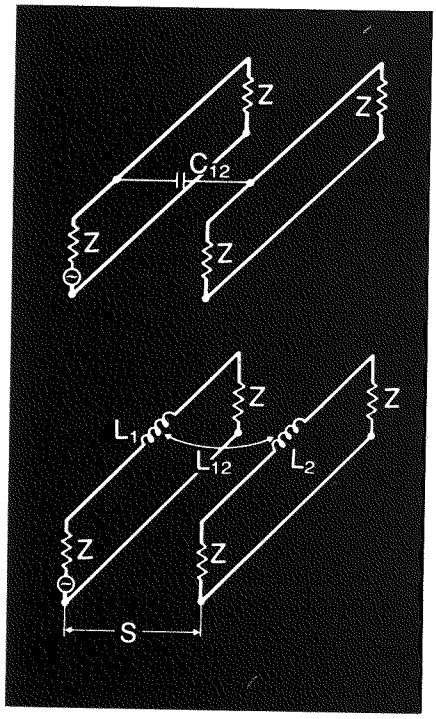


Fig. 9 - Accoppiamento per cross-talk (C.C.C.). C12 e L12 rappresentano rispettivamente, la capacità e l'induttanza di accoppiamento; S è la separazione (in mm) dei conduttori.

4.1 Standard in Europa

Il principale Ente internazionale generatore di "raccomandazioni" è l'IEC (International Electrical Committee), in accordo con vari Enti di governo.

L'IEC comprende tra l'altro il CISPR (Comité International Special des Perturbations Radioelectriques) che si occupa di interferenze generate da sistemi non destinati alle telecomunicazioni, come nel caso EDP.

Le norme CISPR rivestiranno una importanza sempre maggiore in ambito europeo.

Il principale Ente governativo con forza di legge è invece l'FTZ (FornmeldeTechnikZentralamt) della Repubblica Federale Tedesca. Nel suo ambito, il VDE (Verband Deutscher Elektrotechniker) è l'istituto autorizzato per i test di aderenza ai propri Standard e per la certificazione dei prodotti.

Il bollino VDE consente di ottenere il "permesso di funzionamento" dell'apparecchio in territorio tedesco.

4.2 Standard in U.S.A.

Tra gli enti statunitensi generatori di Standards ricordiamo, tra gli altri, l'IEEE (Inst. of Electrical and Electronic Eng. - società professionale), l'EIA (Electronic Industry Ass. - associazione commerciale) e l'FCC

(Federal Communication Committee), Ente governativo i cui Standards hanno forza di legge.

Il coordinamento generale è affidato all'ANSI (American National STDs Institute).

Le normative FCC rivestono un ruolo vincolante per le aziende del settore; un prodotto privo di certificazione FCC non può venire importato negli Stati Uniti e, di conseguenza, non può esservi oggetto di alcuna forma di Marketing.

La normativa FCC cui devono ottemperare i prodotti EDP è la "Part 15-J" suddivisa in due sottoparti:

La Classe A che impone limiti alle E.M.I. irradiate per le apparecchiature da ambiente non residenziali (es. DPS 4000 ecc.)

La Classe B molto più restrittiva, per apparecchiature installabili anche in ambiente residenziale (es. Personal Computer, stampanti ad aghi ecc.)

Una omologazione di Classe B avviene solo dopo presentazione di un Test Report e relativa documentazione di supporto attestante le misure E.M.I. effettuate e la loro aderenza allo Std.

Le misure E.M.I. devono essere effettuate in un sito di rilevamento (detto Free Field) omologato FCC.

4.3 Standards in Italia

Non esiste attualmente in Italia una Normativa E.M.C. avente forza di Legge.

Le Aziende che sviluppano prodotti per il mercato italiano possono seguire la Normativa CISPR 22, che però è operante solo a livello di "Raccomandazione". Tale Normativa si rifà a quella VDE della Repubblica Federale Tedesca.

È comunque scontato che le Aziende che producono per i mercati Esteri devono attenersi alle Normative vigenti nei Paesi di esportazione.

4.4 Standards nel Resto del Mondo

Le normative E.M.C. dei restanti paesi fanno in genere riferimento a quelle citate, o a livello di Raccomandazione oppure incorporandole (totalmente o parzialmente) nelle varie leggi nazionali.

4.5 Free Fields

Le Figg. 10 e 11 mostrano, rispettivamente, lo schema generale di un Free Field e uno dei due impianti di tale tipo operanti presso in Centro di Ricerca e Sviluppo della *Bull Italia* a Pregnana Milanese. Scopo di tali impianti è, come detto in precedenza, raccogliere e quantificare tramite le antenne e gli strumenti previsti l'E,M.I. irradiata dalle apparecchiature sotto test, al fine di consentire ai progettisti il raggiungimento (compliance) degli Standards per i prodotti in sviluppo.

I Free Fields di Pregnana consentono misure a 10 metri e a 30 metri di distanza dall'apparato sotto test. Il sito a 10 metri è omologato FCC.

5. Gli interventi E.M.C.

Il grafico di Fig. 12 evidenzia il legame esistente tra l'efficacia e il "costo" degli interventi E.M.C. e le fasi di sviluppo di un prodotto EDP.

Appare evidente l'esistenza di due fasi principali di intervento, quella a livello di "Progetto" e quella a livello di "Impianto" (in campo).

5.1 Interventi a livello Progetto

Costituiscono l'insieme degli interventi principali e più profittevoli.

Alcune tra le problematiche da affrontare in questa fase possono essere sunteggiate come segue:

- Corretta stesura dei layout dei circuiti stampati (PWB) al fine di evitare soprattutto problemi di tipo C.M.C. e C.C.C.

  (L'uso di PWB multilayer e di sistemi C.A.D./C.A.E. consente di raggiungere ottimi risultati; molta attenzione deve essere prestata nel gestire modifiche manuali alla tracciatura durante i debugging).
- Corretto dimensionamento delle piste di massa, con eventuale duplicazione delle stesse sia per problemi di C.C.C. che di C.G.C.
- Corretto packaging di sistema, per evitare problemi (soprattutto di C.M.C. e D.M.C.) tra piastre e cavi che le collegano.

GLOSSARIO

E. M. I. **Electro Magnetic Interference:** il complesso dei fenomeni di interferenza irradiata e condotta.

E. M. C. **Electro Magnetic Compatibility:** la disciplina che tratta i metodi di prevenzione/riduzione delle interferenze.

C. G. C. **Common Ground Coupling:** interferenza per disturbi che insorgono su tratti di massa comuni a più dispositivi.

C. M. C. **Common Mode Coupling:** interferenza per disturbi causati da onde e.m. incidenti su loop di massa.

D. M. C. **Differential Mode Coupling:** interferenza per disturbi causati da onde e.m. incidenti su cavo di trasmissione segnali.

C. C. C. **Cable to Cable Coupling:** interferenza per disturbi causati per accoppiamento capacitivo o induttivo tra cavi segnale vicini (crosstalk).

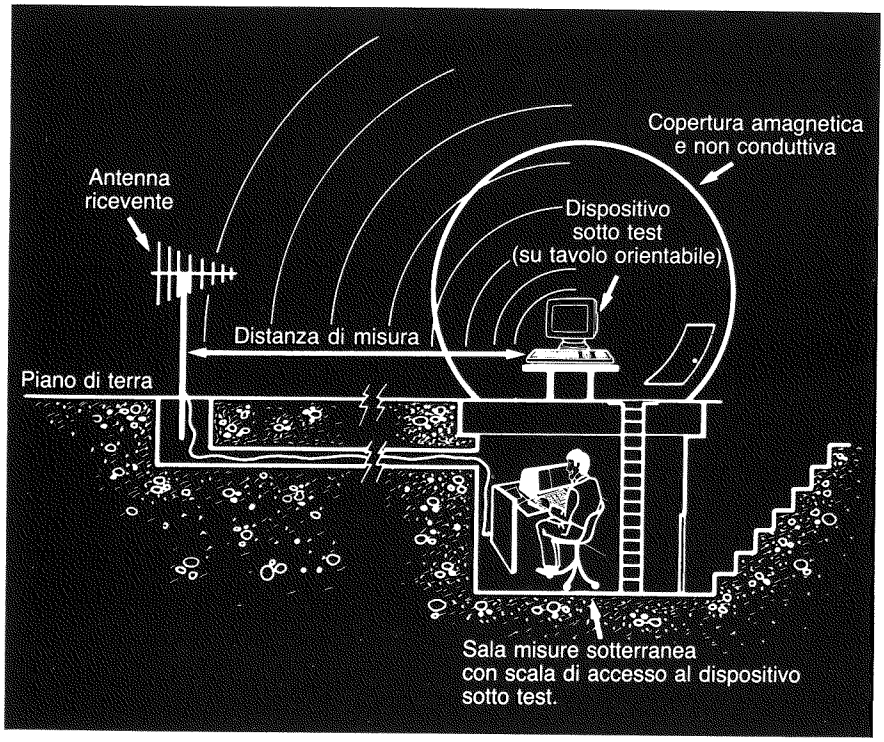


Fig. 10 - Sezione di free-field con sala misure sotterranea.
La sala misure può anche essere a livello suolo e remota.

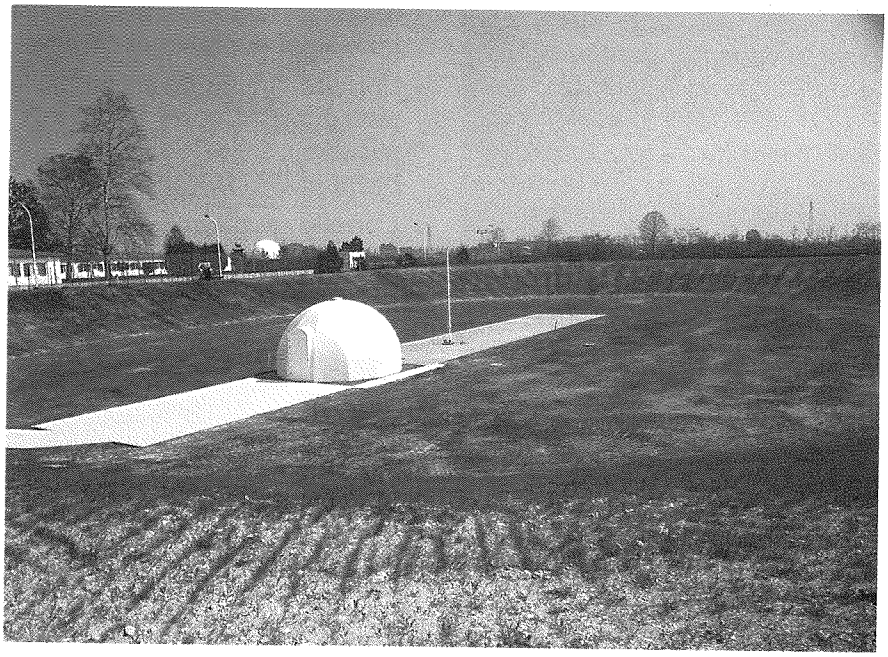


Fig. 11 - Impianto free-field per rilevamenti a 30 m realizzato presso Il Centro di Ricerca e Sviluppo della Bull Italia a Pregnana Milanese. Per migliorare le funzionalità, l'impianto è realizzato al di sotto del piano di campagna, in un sito di forma ellittica appositamente scavato. La centrale di misura e comando è remota.

- Definizione di un "System Ground Point" di sistema, cui riferisce tutti i cavi di massa (giallo-verdi) interni alla macchina, sia riguardanti l'elettronica che il frame meccanico.
- Configurazione dei collegamenti di segnale e di massa dei cavi che entrano/escono dalle varie interfacce dell'unità.

I metodi per affrontare correttamente tali problemi consistono nel condurre, se il caso, analisi preliminari sul contenuto spettrale dei segnali digitali più veloci (fronti e frequenze fondamentali) ed una scelta tra le varie tecniche di riduzione di sensibilità ai disturbi che la teoria propone in funzione del tipo di interferenza prevedibile.

Momento cruciale dello sviluppo in ottica E.M.C. di un nuovo hardware è infine la verifica a livello Prototipo delle E.M.I. irradiate al Free Field, pianificando gli eventuali interventi correttivi sull'hardware dopo un'attenta valutazione dei risultati delle prove.

Fig. 12 - Fasi di sviluppo di un sistema elettronico e interventi EMC.

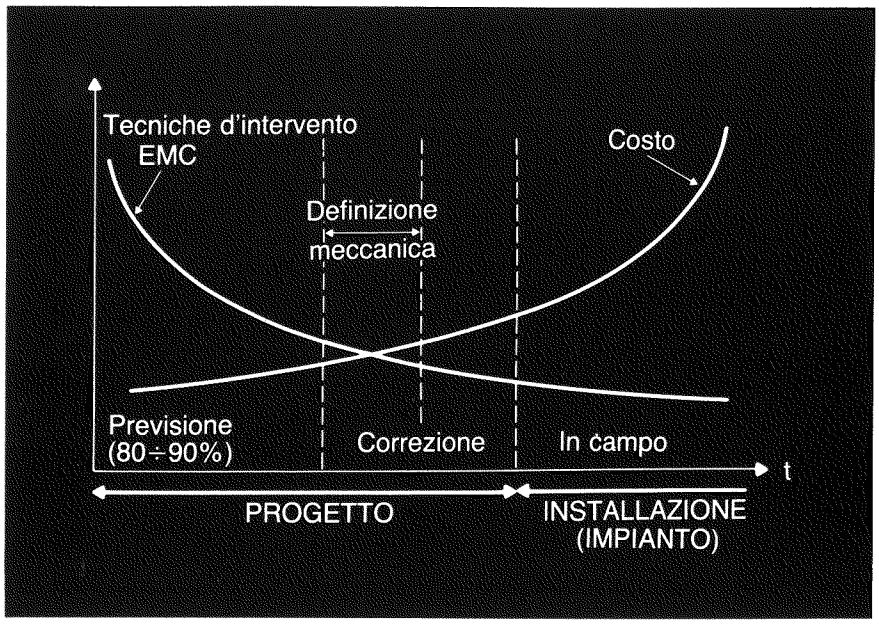
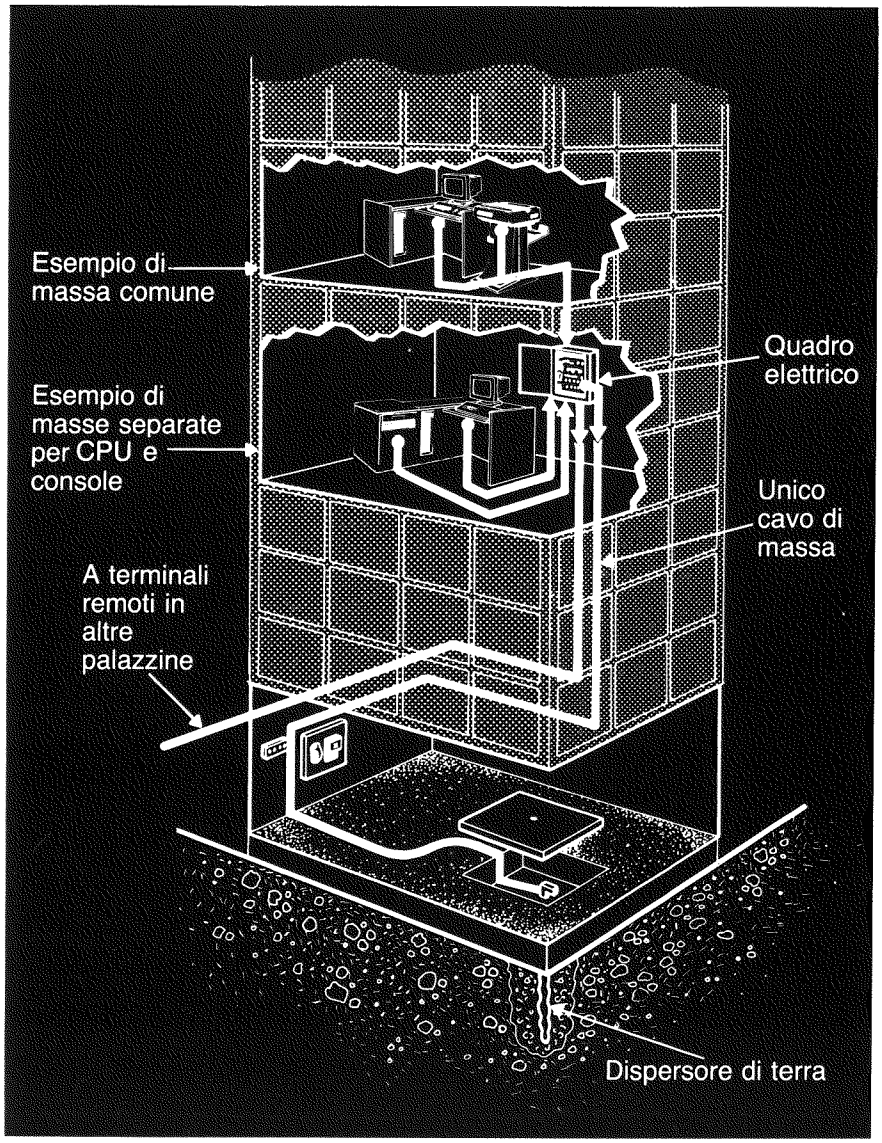


Fig. 13 - Esempio di distribuzione di ground in edificio. Non sono riportati i cavi di alimentazione ed i cavi segnale.



5.2 Interventi a livello Impianto

Lo scopo dichiarato del Progetto in ottica E.M.C. è dunque quello di ottenere un prodotto, sia esso Periferica o Unità Centrale, intrinsecamente conforme alle Norme Internazionali.

Rimane da esaminare il secondo aspetto delle problematiche E.M.C., quello dei problemi di "Interferenza" che possono insorgere in Field a livello "impianti EDP", nelle più varie conditioni di hardware e di "site" di installazione. Tali fenomeni, spesso misconosciuti e/o attribuiti ad altre cause, meritano di essere esaminati più in dettaglio per i loro possibili impatti a livello Servizio Assistenza Clienti.

La connessione hardware che si presenta più frequentemente a livello Sala Calcolo o Personal Computing (ufficio) si può considerare quella rappresentata in Fig. 6a.

Da ciò consegue che fenomeni di tipo C.G.C. e C.M.C. possono insorgere piuttosto facilmente in tali installazioni.

I problemi E.M.I. si verificano laddove i disturbi elettrici superano le soglie di immunità delle varie apparecchiature (videoterminali, P.C., stampanti ecc...).

Serve sicuramente ricordare che alcuni dei cosiddetti "malfunzionanti intermittenti", ben noti in Field, possono essere causati da tali fenomeni e meritano, quindi, una analisi in ottica E.M.I.

Un altro esempio: I riferimenti di ground (Dispersori di terra) in caso di impianti EDP distribuiti tra più palazzine dello stesso complesso di uffici sono, in alcune installazioni, molteplici e distanti tra loro. Ciò costituisce possibile esca a fenomeni E.M.I. determinati da scariche atmosferiche (fulmini) sviluppatesi nelle vicinanze (fino a vari Km.), in virtù dei gradienti di potenziale indotti nel terreno in cui sono infissi i dispersori.

Ciò rivela il ruolo fondamentale che riveste, molto spesso, la "Rete di Grounding"; in Fig. 13 si riporta un classico esempio di sala calcolo/ufficio vista nell'ambito di un intero fabbricato; da notare la corretta distribuzione di masse (senza tratti comuni) per l'unità centrale e la console ed il conduttore di massa destinato ai terminali in altre palazzine che diparte dallo stesso quadro elettrico (unico dispersore).

E quindi molto importante per un tecnico/impiantista saper valutare cosa esiste "oltre la spina di rete" dell'apparato su cui sta effettuando l'intervento, ed in sintesi:

- il tipo di conduttore di ground, il suo percorso e le sue diramazioni/connessioni.
- Il sistema di distribuzione dell'energia elettrica (sia trifase che monofase) che alimenta l'impianto ed il conseguente grounding a livello edificio.
- La disposizione e la funzionalità del sistema di dispersori di terra.
- La reale presenza di un problema E.M.I.
- La procedura più conveniente (impatti tecnici, rapporti con il cliente) per risolvere il problema in oggetto.

6. La diffusione del Know-How

Appare evidente che una più vasta diffusione di conoscenze specifiche sul tema Compatibilità Elettromagnetica è ormai d'obbligo in aziende di informatica come il *gruppo Bull.*

Allo scopo di favorire tale diffusione e nel quadro del progetto Q.I.P. (Quality Improvement Program), sono stati introdotti, su queste tematiche, due Corsi di Formazione Specialistica.

Il loro obiettivo è quello di mettere in grado due precise figure professionali (il Progettista Hardware ed il Tecnico di Field) di progettare/intervenire in ottica E.M.C.

I contenuti riguardano, infatti, le metodologie preventive e correttive da applicare sia nelle fasi di sviluppo di nuovi prodotti che negli interventi sui Sistemi installati in Field in presenza di fenomeni di Interferenza Elettromagnetica.

Bibliografia

[1] E. NANO, *Compatibilità Elettromagnetica,* Quaderni di Elettronica Boringhieri N. 4.

[2] D. WHITE, K. ATKINSON, J. OSBURN, *Taming EMI in microprocessor systems,* IEEE Spectrum, Dec. 1985.

[3] *IEEE Transactions on ELECTROMAGNETIC COMPATIBILITY,* Numeri vari.

[4] C.N.R. - C.E.I., Norme per gli impianti di messa a terra.

[5] L. INZOLI, *Vincoli commerciali dovuti alla Compatibilità Elettromagnetica,* Atti del Convegno BIAS Microelettronica, Milano 1986.

[6] S. CANIGGIA, L. INZOLI, *Approccio alla Compatibilità Elettromagnetica per Sistemi Elettronici,* FAST, 1986.

Conversazioni col computer

FRANCO FILIPPAZZI

Bull Italia
Centro Studi Informatica e Automazione
Milano

1. Introduzione

Circa 25 anni or sono, Joseph Weizenbaum, docente di computer science al celebre MIT di Boston, realizzava un programma di nome ELIZA.

È questo il nome della protagonista del "Pigmalione" di G.B. Shaw; nella famosa commedia, l'incolta e rozza fioraia Eliza impara ad esprimersi in modo forbito, così che, alla fine, nessuno è in grado di immaginare le sue vere origini.

Analogamente, il programma ideato da Weizenbaum consente ad un computer di sostenere una conversazione in modo tale da poter essere scambiato per una persona.

A distanza di anni, si può ancora apprezzare la genialità di questo programma e il suo intento provocatorio nei riguardi dell'intelligenza delle macchine.

Riteniamo perciò utile riproporre qui ELIZA, certamente uno dei programmi più famosi della storia dell'informatica.

2. Conversare con una macchina

È possibile che un computer sostenga una conversazione con un interlocutore umano riuscendo a mascherare la propria natura?

In altri termini, può una macchina dare risposte plausibili, pur senza capire ciò di cui si parla?

C'è da dire che il rapporto tra capire e rispondere in modo plausibile non è netto come potrebbe sembrare. Si pensi, ad esempio, ad una conversazione tra più persone, una delle quali non è interessata al discorso ma non vuol darlo a vedere. Essa può celare la sua mancanza di attenzione interloquendo di tanto in tanto con frasi non compromettenti, quali "interessante...", "veramente?...", o ripetendo meccanicamente parti di frasi pronunciate da altri. Con una certa abilità, questa tecnica può essere applicata con successo, anche se ovviamente ha dei limiti.

In effetti, perchè un computer possa sostenere una conversazione senza smascherarsi come tale dopo poche batture, occorre che soddisfi tre requisiti generali.

Il primo è che ci si possa rivolgere ad esso in linguaggio naturale, nello stesso modo cioè con cui si parlerebbe ad una persona, senza limitazioni lessicali o sintattiche.

Il secondo è che gli interventi del computer siano formalmente coerenti col tema della conversazione.

Il terzo è che il computer dia risposte equivalenti a fronte di una medesima domanda, anche se questa è formulata in modo diverso.

Per realizzare l'ambizioso obiettivo di far conversare un computer in modo plausibile, è ovviamente opportuno delimitare l'ambito del dialogo.

Nel suo esperimento, Weizenbaum scelse come ambito lo studio di uno psicoterapeuta, in cui il dottore (di scuola rogeriana) è il computer, ossia ELIZA, e il paziente è l'uomo.

Questa scelta di Weizenbaum non era priva di malizia; c'era, infatti, da parte sua l'intenzione di prendere in giro certe pratiche psicologiche, mostrando come il dottore potesse essere sostituito da una macchina.

In effetti, il programma costruito da Weizenbaum risultò così ben congegnato da dare l'impressione che la macchina capisse realmente ciò che diceva il suo interlocutore. Per rendersene conto, si veda nel riquadro (Fig. 1) un brano estratto da una effettiva conversazione con ELIZA. A prima vista, nessuno direbbe che uno dei due interlocutori è un computer.

Ma come è possibile che una macchina possa mascherare tanto bene la propria identità?

3. Un meccanismo semplice ma efficace

Il procedimento con cui ELIZA costruisce le sue risposte si può riassumere nel modo seguente (Fig. 2 e 3).

1) Ogni frase dell'interlocutore umano (il "paziente") viene analizzata per cercarvi una parola (o locuzione) chiave. Un certo numero di tali parole, opportunamente scelte, è previsto nel programma. Possono evidentemente verificarsi due possibilità: la frase del paziente contiene una o più parole chiave, oppure non ne contiene nessuna.

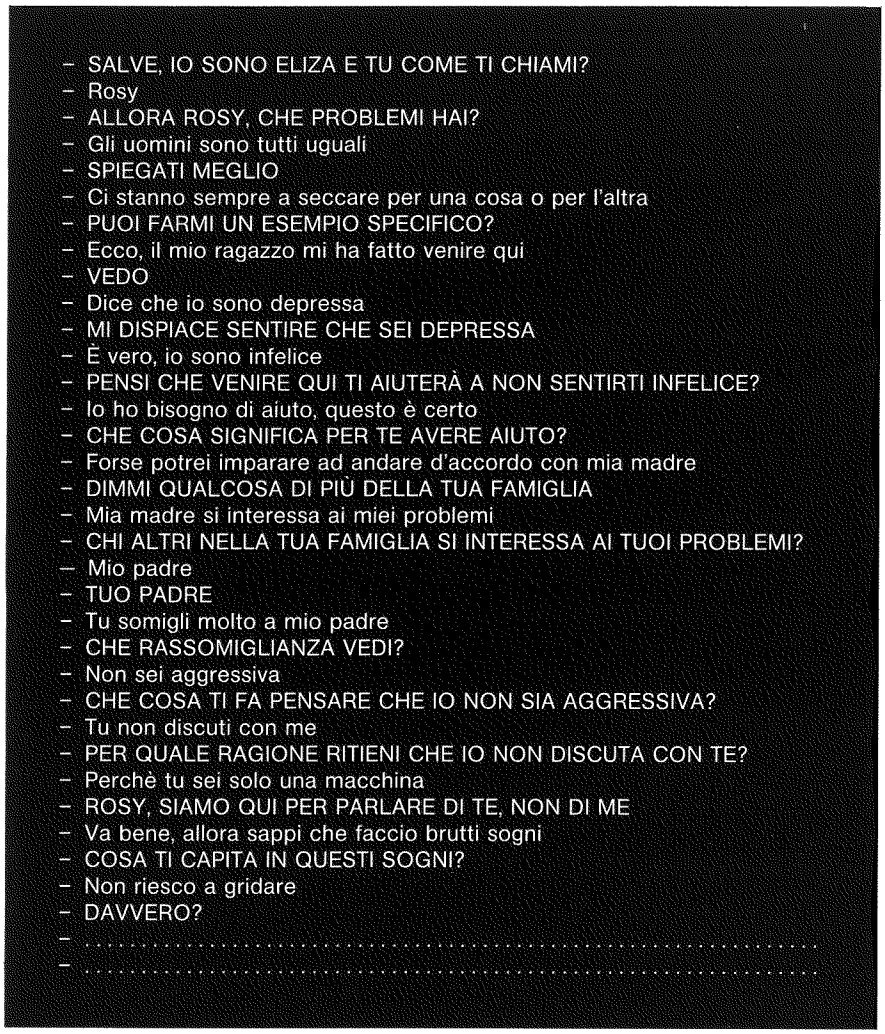


Fig. 1 - Brano estratto da una effettiva conversazione con ELIZA. (In maiuscolo le frasi della macchina, in minuscolo quelle della persona).

2) Nel primo caso, il programma seleziona una delle parole chiave (in base a criteri fissati di priorità) e associa a tale parola una risposta preconfezionata. In effetti, per ogni parola chiave non è prevista una sola risposta, ma un gruppo di risposte *equivalenti;* tra queste il programma estrae via via la risposta da dare.

3) Un ulteriore (ed efficacissimo) accorgimento consiste nell'inglobare nella risposta una porzione della frase del paziente. Ciò richiede che il programma sezioni secondo certi criteri la frase sottopostagli, ne estragga una parte e la coniughi opportunamente prima di "montarla" nella frase di risposta.

4) Nel caso non ci sia alcuna parola chiave nella frase del paziente, viene data una risposta interlocutoria, del tipo: "Spiegati meglio", "Davvero?", "Prosegui", ecc. Risposte di questo tipo possono sempre entrare in una conversazione, purché non troppo spesso. Perciò, di tanto in tanto, il programma introduce qualcosa di diverso e cioè una espressione del tipo "Prima tu hai detto che..." seguita da una frase pronunciata in precedenza dal paziente.

Per meglio capire il meccanismo di "smontaggio" e "rimontaggio" ora descritto, facciamo un esempio.

Si prenda in esame la frase: *"Io sono molto preoccupato in questi giorni".* Supponiamo che uno straniero con una limitata conoscenza dell'italiano ma con un buon orecchio, abbia sentito pronunciare questa frase ma abbia capito solo le prime due parole (*Io sono*). Desiderando sembrare interessato, egli può rispondere: "*Perchè lei è molto preoccupato in questi giorni?*"

Ciò che egli ha fatto è stato di applicare alla frase originaria una "maschera", una parte della quale coincide con le parole "*Io sono*", mentre l'altra isola ciò che viene dopo.

Egli deve avere anche un criterio di rimontaggio associato specificatamente a questa maschera, il quale prescrive che ogni frase della forma "*Io sono BLABLA*" venga trasformata in "*Perchè lei è BLABLA?*", qualunque sia il significato di BLABLA.

Un esempio un po' più complicato è dato dalla frase: "*Tutto sommato, credo che tu mi sia simpatico*". Ora lo straniero afferra la locuzione "*credo che*", nonchè il fatto che nel BLABLA che segue ci sono dei pronomi e dei verbi. Egli applica una maschera che scompone la frase ancora in due parti:

1) *Tutto sommato, credo che*

2) *tu mi sia simpatico*

Fig. 2 - Meccanismo di "smontaggio" di una frase e di "montaggio" della risposta.

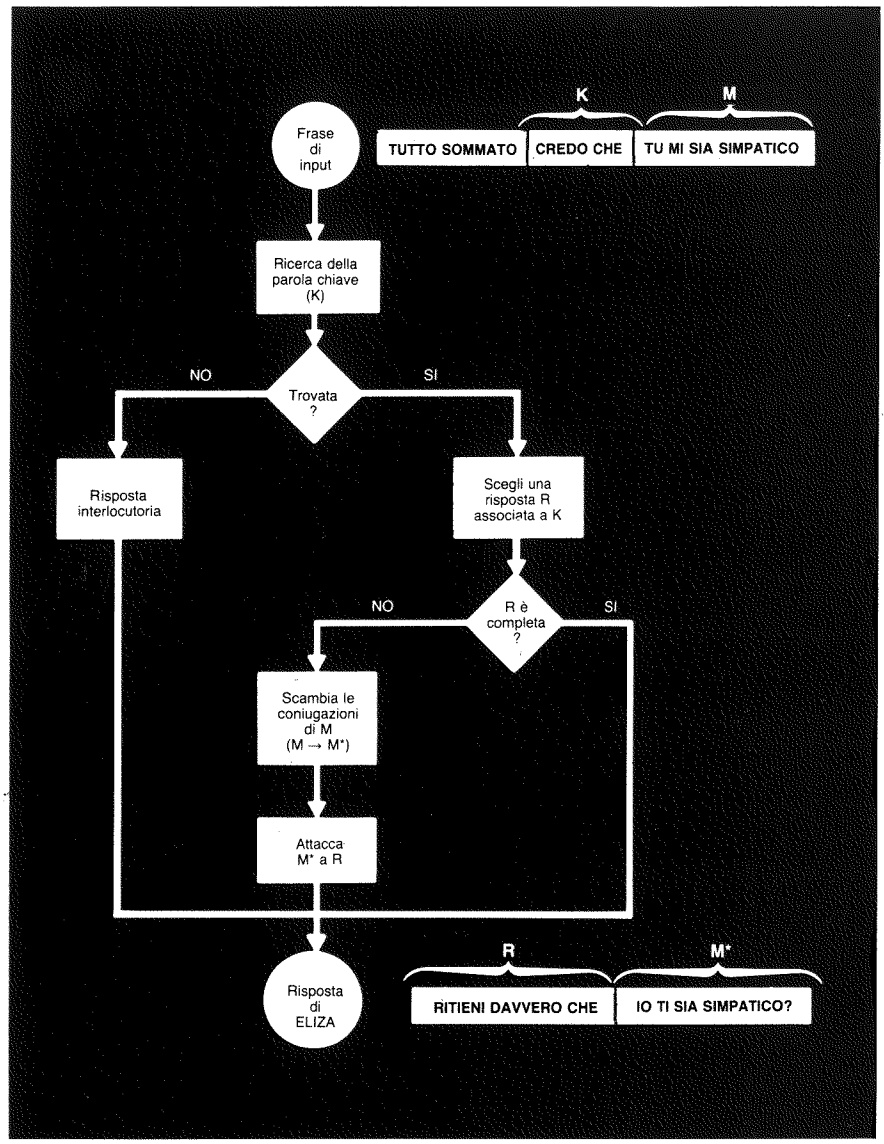
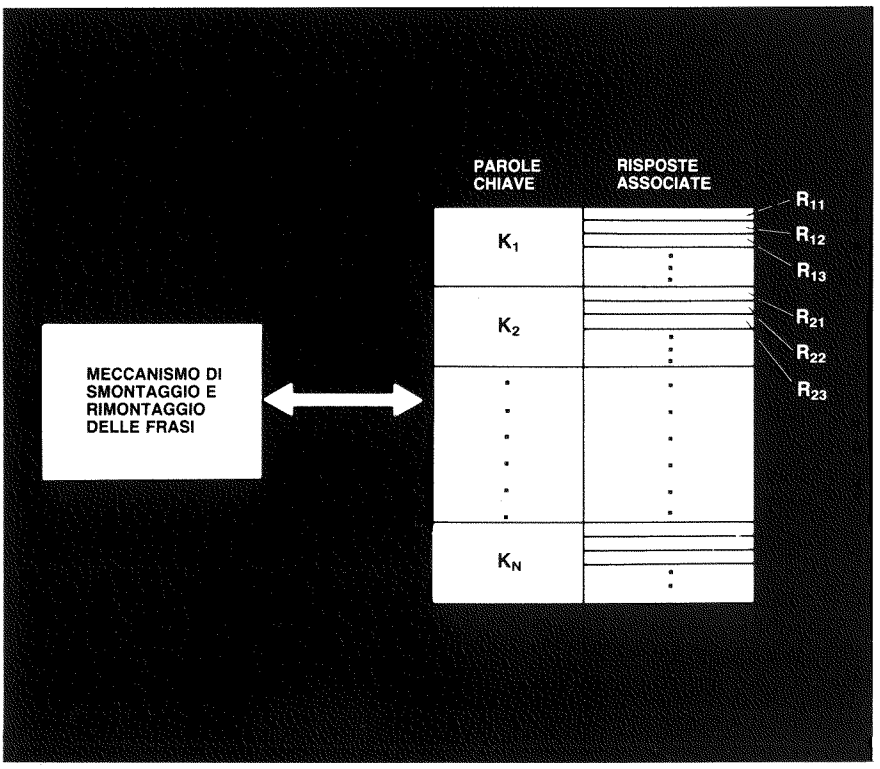


Fig. 3 - ELIZA risulta, in sostanza, costituita da due parti, di cui una è invariante, mentre l'altra (parole chiave e risposte associate) è fortemente correlata all'ambito in cui si svolge la conversazione. Cambiando questa seconda parte, la macchina assume ruoli e personalità diversi.



La regola di rimontaggio può essere allora la seguente:

* associare alla parte 1) una frase di uso generale come, ad esempio, "*Ritieni davvero che*";
* coniugare a rovescio pronomi e verbi contenuti nella parte 2);
* attaccare la parte 2) così modificata, con un punto interrogativo finale.

La risposta diventa quindi: "*Ritieni davvero che io ti sia simpatico?*"

Come si vede, una risposta perfettamento a tono, anche se lo straniero non ha, in effetti, capito per niente il significato della frase rivoltagli.

Un meccanismo di questo tipo stà alla base delle risposte di ELIZA.

Ovviamente la sua efficacia dipende da vari fattori: dalle regole di scomposizione/ricomposizione delle frasi, dalla scelta delle parole chiave, dal repertorio delle frasi di risposta, dalla completezza delle coniugazioni previste.

È veramente sorprendente come attraverso un meccanismo così semplice (si pensi che l'intero programma ELIZA consta di poche centinaia di righe), una macchina possa tenere conversazioni naturali e credibili, come quelle riportate nel riquadro.

4. Esperienze con ELIZA

Il programma di Weizenbaum venne fatto circolare in diverse università americane, suscitando non di rado reazioni sconcertanti, come lo stesso Weizenbaum ebbe poi a raccontare nel suo libro "*Computer power and human reason*" [2].

Egli rimase colpito da come persone non sprovvedute si lasciassero coinvolgere dalle conversazioni con ELIZA fino al punto da scordare la sua natura e attribuirle una vera e propria personalità

Weizenbaum racconta, ad esempio, come la sua segretaria, pur sapendo che si trattava solo di un programma, cominciò ad avere con esso un rapporto emotivo, al punto da chiedere di essere lasciata sola nella stanza del computer per parlare con ELIZA di argomenti strettamente privati.

Weizenbaum scoprì, inoltre, che ELIZA veniva consultata nottetempo da molte persone usando i terminali distribuiti nel campus. Egli ricevette anche delle telefonate da gente che si lamentava delle lunghe attese per accedere al programma, avendo urgenti questioni personali da porre.

Un altro fatto che lo lasciò sconcertato fu che alcuni docenti di psichiatria attribuissero al programma un effettivo valore terapeutico, al punto da proporre che una versione estesa e perfezionata fosse messa a disposizione dei pazienti nelle cliniche per malattie mentali.

A prescindere da facili ironie su certe pratiche psicoterapeutiche, ELIZA offre lo spunto a riflessioni su che cosa realmente debba intendersi per "intelligenza".

Non è negli scopi di questo articolo dibattere un tema così impegnativo e opinabile. Ci limitiamo qui soltanto a far notare alcune sottili provocazioni insite nel lavoro di Weizenbaum.

ELIZA sembra capire, ma in realtà non capisce affatto. Le sue risposte sono il risultato di un mero automatismo che associa parole chiave a frasi preconfezionate.

Ma quante conversazioni tra persone non sono, in fondo,riconducibili a schemi di questo tipo?

I discorsi di ELIZA sono formalmente corretti; però, a ben vedere, essa non affronta mai l'argomento, è evasiva, ripropone domande a lei rivolte, divaga.

Anche qui è fin troppo facile osservare che questa abilità non è peculiare della macchina; quante volte ci è capitato di sentire persone fare lunghi discorsi senza, in realtà, dire niente?

Infine, il successo di ELIZA presso i "pazienti" non è che la conferma del fatto che abbiamo bisogno di qualcuno che ci stia ad ascoltare. Fosse anche una macchina...

Bibliografia

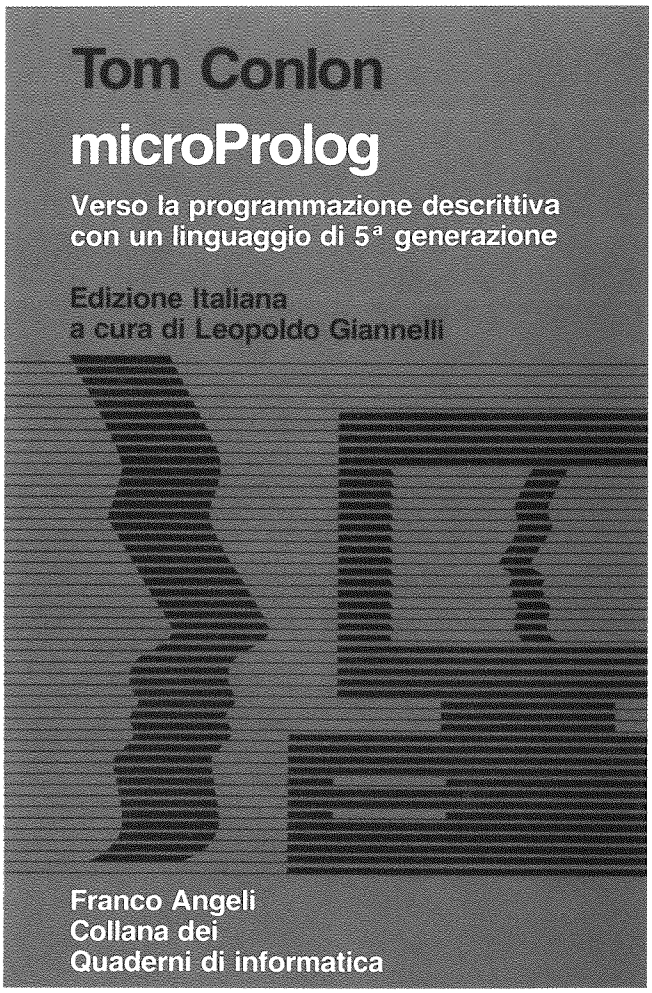
[1] J. WEIZENBAUM, *ELIZA, a computer program for the study of natural language communication between man and machine, Comm. of ACM, Vol. 9, N. 1 (1966).*

[2] J. WEIZENBAUM, *Computer power and human reason,* Freeman and Co. (1976).

BIBLIOTECA NOVITÀ

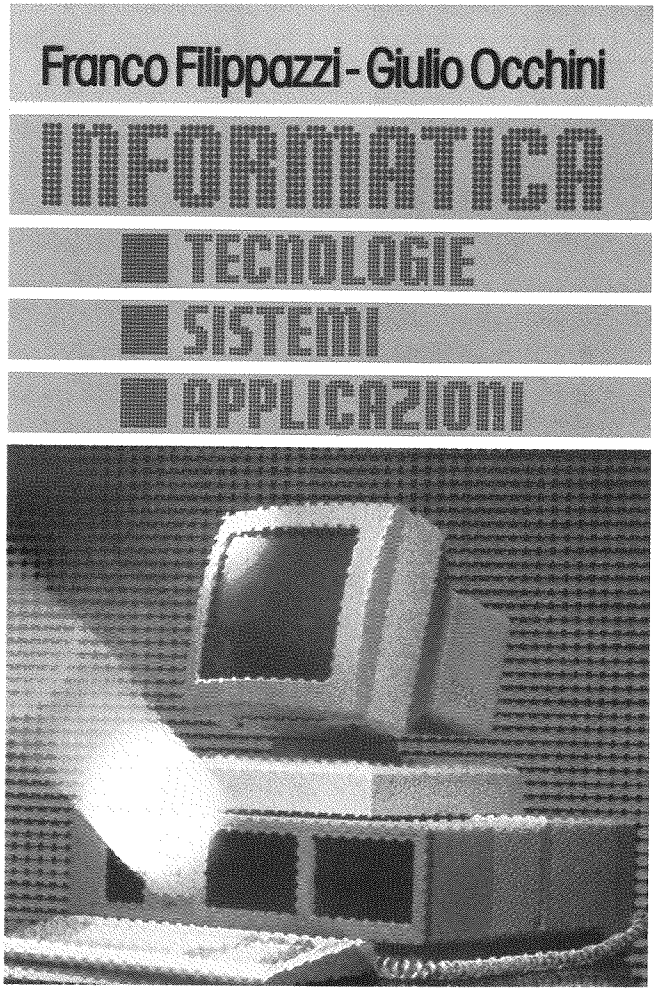
I TESTI DEI QUADERNI DI INFORMNATICA

Tom Conlon: Micro Prolog - Verso la programmazione descrittiva con un linguaggio di 5ª generazione
Ediz. Italiana a cura di Leopoldo Giannelli. Collana dei Quaderni di Informatica, Ed. F. Angeli, 1989 (pag. 210, L. 24.000)



PROLOG (abbreviazione di PROgramming in LOGic) è il linguaggio dichiarativo di maggior successo sviluppato ad oggi.
Adottato dai giapponesi come linguaggio-«nocciolo» per il loro progetto di V generazione, PROLOG è gia stato usato per lo sviluppo di sistemi esperti, per l'analisi di linguaggi naturali, per dimostrare teoremi matematici, per generare traslatori per computer, per risolvere problemi. PROLOG è basato principalmente su parole del linguaggio corrente - non su simboli oscuri - e impiega queste parole per descrivere fatti, relazioni e schemi in modo logico, conciso.
Oggi, Micro-PROLOG, una significativa e promettente implementazione di PROLOG, è disponibile per i PC basati su MS DOS e per la maggior parte dei sistemi CP/M e stà per essere rilasciato anche per numerosi altri computer, tra i quali Apple II. Con questo volume programmatori, studenti, insegnanti e appassionati dei PC possono iniziare ad esplorare il mondo di PROLOG scrivendo e facendo girare programmi.
Tom Conlon ci dà, infatti, un'indispensabile introduzione informale a questo linguaggio rivoluzionario, ponendo l'accento sulla sua logica e sullo sviluppo delle abilità necessarie per diventare dei bravi *problemsolver*. Il volume non richiede alcuna conoscenza preliminare di matematica e di programmazione. Scritto in modo vivace ed accessibile, consente d'acquistare una completa padronanza delle basi e delle possibilità di Micro-PROLOG. La trattazione è integrata da esempi e da esercizi proposti al lettore, oltre che da indicazioni utili per ulteriori approfondimenti. Il volume - curato nell'edizione italiana da Leopoldo Giannelli della Bull Italia - è così il testo introduttivo ideale consigliabile come prima lettura a quanti desiderano conoscere questo importante linguaggio.

Franco Filippazzi, Giulio Occhini
Informatica: Tecnologie, Sistemi, Applicazioni
Edizione CDE, 1989 (pag. 430).



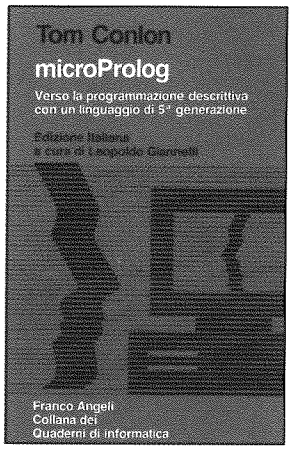
L'informatica è senza dubbio uno degli aspetti che maggiormente caratterizzano l'epoca in cui viviamo. Essa suscita un diffuso interesse per almeno un paio di motivi di fondo.
il primo è la sua "peculiarità". Il computer, infatti, è una macchina che si differenzia da tutte le altre perchè opera in una dimensione - il trattamento dei simboli e della conoscenza - che attiene alla sfera intellettiva dell'uomo.
Il secondo motivo è la sua "pervasività". Non esiste, infatti, settore delle attività umane in cui i metodi, le tecniche e gli strumenti informatici non vengano oggi applicati, o non possano esserlo in futuro.
Con quest'opera, gli autori offrono la possibilità al lettore di entrare nel mondo dei computer attraverso una trattazione di taglio divulgativo, ma organica e rigorosa, accessibile al pubblico non specialista.

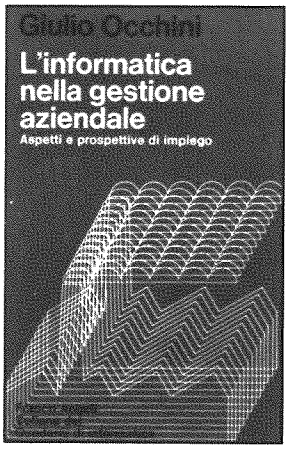
Gli autori sono due noti esperti del settore, responsabili delle attività del «Centro Studi Informatica e Automazione», una struttura della Bull Italia creata per condurre o coordinare ricerche sugli orientamenti di fondo dell'informatica e della automazione.

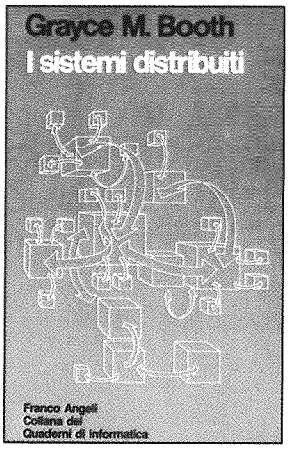
Franco Filippazzi, laureato in fisica, è stato uno dei progettisti del primo elaboratore elettronico italiano («Elea»). Ha dato contributi originali alla tecnologia dell'elaboratore, descritti in numerosi brevetti e pubblicazioni. Parallelamente all'attività industriale, svolge un ruolo attivo in ambito culturale, sia con l'insegnamento universitario sia con responsabilità in associazioni scientifico-tecniche.

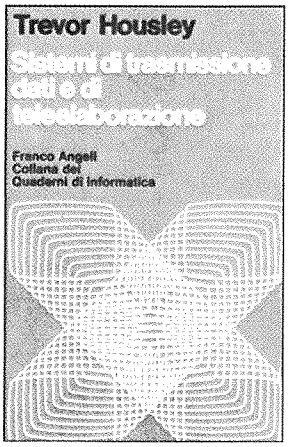
Giulio Occhini, laureato in fisica, vanta una esperienza di oltre 25 anni nel campo delle applicazioni informatiche. È autore di numerosi contributi sugli aspetti economico-organizzativi del processo di automazione ed è presente, con funzioni direttive, nelle principali associazioni scientifiche del settore.

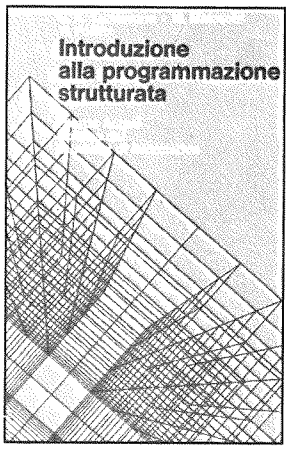
I TESTI DEI
QUADERNI DI INFORMATICA

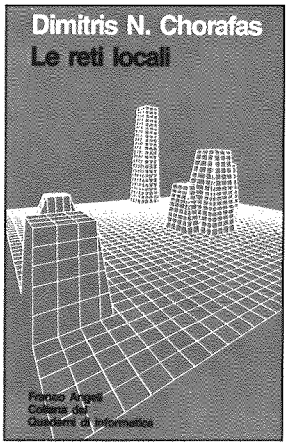












Spedizione in abbonamento postale, Gruppo IV, 2° semestre.
Pubblicità Inferiore al 70%.
N. 36, Maggio - Luglio 1989
TAXE PERÇUE

P.P. Tassa riscossa PT Piacenza F. Italia